

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:28:34 ; Search time 0.810811 Seconds
(without alignments)
830.671 Million cell updates/sec

Title: us-10-820-200-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79:**

2: PIR1:**

3: PIR2:**

4: PIR3:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	478	2 JK0201	alpha-amylase (EC
2	40	100.0	498	2 A48305	alpha-amylase (EC
3	40	100.0	499	1 ALA81	alpha-amylase (EC
4	40	100.0	499	1 ALA81	alpha-amylase (EC
5	40	100.0	499	2 JS0663	alpha-amylase (EC
6	40	100.0	499	2 JT0466	alpha-amylase (EC
7	40	100.0	499	2 JN0588	alpha-amylase (EC
8	40	100.0	499	2 B48305	alpha-amylase (EC
9	40	100.0	624	1 JC4510	pullulanase (EC 3.
10	34	85.0	130	2 H71415	hypothetical prote
11	33	82.5	420	2 T42616	envelope protein -
12	33	82.5	601	2 T18787	hypothetical prote
13	33	82.5	1131	2 T41943	major DNA binding
14	33	82.5	4351	2 T00252	MEGF1 protein - ra
15	32	80.0	144	2 JC7121	androgenic gland h
16	32	80.0	180	2 B95039	hypothetical prote
17	32	80.0	289	2 A43562	homeotic protein H
18	32	80.0	317	2 E84088	restriction endonu
19	32	80.0	351	2 E86187	YUP812.10 [import
20	32	80.0	411	2 S61245	probable virion gl
21	32	80.0	419	2 T15199	hypothetical prote
22	32	80.0	482	2 A90248	conserved hypotet
23	32	80.0	685	2 S67146	probable membrane
24	32	80.0	820	1 DEECK	thrA bifunctional
25	32	80.0	820	2 B85480	aspartokinase I, h
26	32	80.0	820	2 B90629	aspartokinase I-ho
27	32	80.0	820	2 AC0502	aspartokinase I/h
28	31	77.5	67	2 T17658	hypothetical prote
29	31	77.5	104	2 S49803	hypothetical prote

30	31	77.5	214	2 B75513	pyridoxamine 5-pho
31	31	77.5	261	2 AF1307	hypothetical prote
32	31	77.5	261	2 AF1679	hypothetical prote
33	31	77.5	386	2 S72435	RNA-binding protei
34	31	77.5	461	2 F86597	UDP-glucose pyroph
35	31	77.5	461	2 H72024	UDP-N-acetylglucos
36	31	77.5	461	2 T35746	hypothetical prote
37	31	77.5	614	1 I48385	RNA helicase TN22
38	31	77.5	614	2 J101087	RNA helicase, ATP-
39	31	77.5	668	2 T10575	hypothetical prote
40	31	77.5	825	2 S55060	fertilin alpha-II
41	31	77.5	833	2 T03864	hypothetical prote
42	31	77.5	905	2 S55059	fertilin alpha-I -
43	30	75.0	182	2 AE2146	hypothetical prote
44	30	75.0	194	2 T22209	hypothetical prote
45	30	75.0	229	2 A89473	protein F52D2.2 [i

ALIGNMENTS

RESULT 1

JK0201

alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae

N;Alternate names: glycogenase; taka-amylase A

C;Species: Aspergillus oryzae

C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004

C;Accession: JK0201

R;Toda, H.; Kondo, K.; Narita, K.

Proc. Jpn. Acad. 58B, 208-212, 1982

A;Title: The complete amino acid sequence of Taka-amylase A.

A;Reference number: JK0201

A;Accession: JK0201

A;Molecule type: protein

A;Residues: 1-478 <TOD>

A;Cross-references: UNIPROT:P10529

C;Comment: One atom of calcium per molecule is essential for the activity.

C;Comment: This enzyme is a glycoprotein.

C;Comment: See also PIR:JT0466 and PIR:JS0240.

C;Function:

A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A;Pathway: glycogen/starch degradation

C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F;173-300/Domain: alpha-amylase core homology <AMY>

F;197/Binding site: carboxylate (Asn) (covalent) #status experimental

F;210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 40; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7

Db 141 SSQDYFH 147

RESULT 2

A48305

alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori

C;Species: Aspergillus awamori

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C;Accession: A48305

R;Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.; J

Curr. Genet. 17, 203-212, 1990

A;Title: Cloning, characterization, and expression of two alpha-amylase genes from Asperg

A;Reference number: A48305; PMID:90254827; PMID:2340591

A;Accession: A48305

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-498 <KOR>

A;Cross-references: UNIPROT:Q02905

C;Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 40; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
 |||||

Db 162 SSQDYFH 168

RESULT 3

ALAS1
 Alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 09-Jul-2004
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
 R:Wirsael, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon o
 A:Reference number: S04548; MUID:89237897; PMID:2785629

A:Accession: S04548

A:Molecule type: DNA

A:Residues: 1-499 <WIR>

A:Cross-references: UNIPROT:P10529; EMBL:X12725; NID:G2430; PIDN:CAA31218.1; PID:G295921

A:Genetics: AMY1

A:Accession: A33214

A:Molecule type: mRNA

A:Residues: 1-499 <W12>

A:Cross-references: GB:X12725; NID:G2430; PIDN:CAA31218.1; PID:G295921

R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A:Reference number: J50240; MUID:89378767; PMID:2789162

A:Accession: J50240

A:Molecule type: DNA

A:Residues: 1-499 <GEN>

A:Genetics: AMY2

A:Note: the authors refer to this as isozyme II

R:Isemura, S.; Ikenaka, T.

J. Biochem. 74, 1-10, 1973

A:Reference number: A91930; MUID:74001521; PMID:4733850

A:Accession: A91930

A:Molecule type: protein

A:Residues: 206-225 <ISE>

R:Narita, K.

Proc. Jpn. Acad. 51, 285-290, 1975

A:Reference number: A93767

A:Accession: A93767

A:Molecule type: protein

A:Residues: 434-443, 446-447, 'O', 449-458, 'GTTV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>

R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370; PMID:6609921

A:Contents: annotation; X-ray crystallography, 3.0 angstroms

R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.

submitted to the Brookhaven Protein Data Bank, August 1992

A:Reference number: A51548; PDB:6TAA

A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497

C:Comment: One atom of calcium per molecule is essential for activity.

A:Genetics: <AMY1>

C:Gene: amy1

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Genetics: <AMY2>

A:Gene: amy2; AmyII

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polysa
 F:1-21/Domain: signal sequence #status experimental <SIG>
 F:22-499/Product: alpha-amylase 1 #status experimental <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
 |||||

Db 162 SSQDYFH 168

RESULT 4

ALAS3

Alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae

C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 09-Jul-2004

C:Accession: S04549; A33215; A44713

R:Wirsael, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.

Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon o

A:Reference number: S04548; MUID:89237897; PMID:2785629

A:Accession: S04549

A:Molecule type: DNA

A:Residues: 1-499 <WIR>

A:Cross-references: UNIPROT:Q96TH4; EMBL:X12727; NID:G2454; PIDN:CAA31220.1; PID:G295922

A:Accession: A33215

A:Molecule type: mRNA

A:Residues: 1-499 <W12>

A:Cross-references: GB:X12727; NID:G2454; PIDN:CAA31220.1; PID:G295922

R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A:Reference number: J50240; MUID:89378767; PMID:2789162

A:Accession: A44713

A:Molecule type: DNA

A:Residues: 1-499 <GEN>

A:Note: the authors refer to this as isozyme I

R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370; PMID:6609921

A:Contents: annotation; X-ray crystallography, 3.0 angstroms

C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics:

A:Gene: amy3; AmyI

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; metal

F:1-21/Domain: signal sequence #status experimental <SIG>

F:22-499/Product: alpha-amylase 3 #status experimental <MAT>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental

F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 40; DB 1; Length 499;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|||||
Db 162 SSQDYFH 168

RESULT 5

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
N;Alternate names: Aspergillus sp.
C;Species: Aspergillus sp.
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C;Accession: J05663
R;Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biosci. Biotechnol. Biochem. 56, 174-179, 1992
A;Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousanii and its expression
A;Reference number: J05663; MUID:92323146; PMID:1368777
A;Accession: J05663
A;Molecule type: mRNA
A;Residues: 1-499 <SHI>
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-499/Product: alpha-amylase #status predicted <ALP>
F;194-321/Domain: alpha-amylase core homology <AMY>
F;218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|||||
Db 162 SSQDYFH 168

RESULT 6

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N;Alternate names: glycogenase; Taka-amylase A
C;Species: Aspergillus oryzae
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: J0466
R;Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A;Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
A;Reference number: J0466
A;Accession: J0466
A;Molecule type: DNA
A;Residues: 1-499 <TAD>
A;Cross-references: UNIPROT:P10529
C;Comment: See also PIR:JK0201 and PIR:J0240.
C;Comment: One atom of calcium per molecule is essential for activity.
C;Genetics:
A;Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-499/Product: alpha-amylase #status predicted <MAT>
F;194-321/Domain: alpha-amylase core homology <AMY>
F;218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|||||
Db 162 SSQDYFH 168

QY 1 SSQDYFH 7
|||||
Db 162 SSQDYFH 168

RESULT 7

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N;Alternate names: Taka-amylase A
C;Species: Aspergillus oryzae
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: J0588
R;Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kiritu, N.; Tsuboi, A.; Uda, S.
Gene 84, 319-327, 1989
A;Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for mul
A;Reference number: J0588; MUID:90128276; PMID:2612911
A;Accession: J0588
A;Molecule type: mRNA
A;Residues: 1-499 <TSU>
A;Cross-references: UNIPROT:Q96TH4
C;Comment: The alpha amylases are encoded by multigene family.
C;Genetics:
A;Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-499/Product: alpha-amylase #status predicted <MAT>
F;194-321/Domain: alpha-amylase core homology <AMY>
F;218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|||||
Db 162 SSQDYFH 168

RESULT 8

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
C;Species: Aspergillus awamori
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: B48305
R;Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.; J
Curr. Genet. 17, 203-212, 1990
A;Title: Cloning, characterization, and expression of two alpha-amylase genes from Asperg
A;Reference number: A48305; MUID:90254827; PMID:2340591
A;Accession: B48305
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499 <KOR>
A;Cross-references: UNIPROT:Q02906
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|||||
Db 162 SSQDYFH 168

```

RESULT 9
JC4510
pullulanase (EC 3.2.1.41) precursor - yeast (lipomyces kononenkoae)
N:Alternate names: LKAI protein; raw starch-degrading amylase
N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
C:Species: Lipomyces kononenkoae
C:Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C:Accession: JC4510; PC4116
R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.
Gene 166, 65-71, 1995
A:Title: Cloning, sequencing, expression analysis and expression in yeasts of a cDNA containing a Lipid
A:Reference number: JC4510; MUID:96105202; PMID:8529895
A:Accession: JC4510
A:Molecule type: mRNA
A:Residues: 1-624 <STE>
A:Cross-references: UNIPROT:Q01117; GB:U30376; NID:g1173536; PIDN:AAC49622.1; PID:g11735
A:Experimental source: strain IGC4052B
A:Accession: PC4116
A:Molecule type: protein
A:Residues: 29-44 <ST2>
A:Experimental source: IGC4052B
C:Genetics:
A:Gene: LKAI
C:Function:
A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
A:Pathway: glycogen/starch degradation
C:Superfamily: Lipomyces alpha-amylase; alpha-amylase core homology; glucoamylase starch
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-624/Product: alpha-amylase #status predicted <MAT>
F:48-141/Domain: glucoamylase starch-binding domain homology <SBD>
F:320-447/Domain: alpha-amylase core homology <AMY>
F:177-185,297-311,397-430,587-622/Disulfide bonds: #status predicted
F:304,344/Binding site: carbohydrate (Aen) (covalent) #status predicted
F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 288 SSQDYFH 294

RESULT 10
H71415
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C:Accession: H71415
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutte, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
Avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: H71415
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <BEV>
A:Cross-references: UNIPROT:O23373; GB:Z97338; NID:G2244870; PID:e326912; PID:G2244877
C:Genetics:
A:Map position: 4C09-4G3845

Query Match 85.0%; Score 34; DB 2; Length 130;
Best Local Similarity 85.7%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 SSQDYFH 7
DB 38 SSADYFH 44

RESULT 11
T42616
envelope protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42616
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A:Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: Z22173; MUID:98264497; PMID:9603335
A:Accession: T42616
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-420 <TEL>
A:Cross-references: UNIPROT:O39309; EMBL:AF030027; NID:g2605950; PIDN:AAC59593.1; PID:g26
A:Experimental source: strain NS80567
C:Genetics:
A:Note: 73
C:Superfamily: pseudorabies virus glycoprotein gp63

Query Match 82.5%; Score 33; DB 2; Length 420;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SQDYFH 7
DB 114 NQDYFH 119

RESULT 12
T18787
hypothetical protein BE10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18787
R:McMurray, A.
Submitted to the EMBL Data Library, March 1997
A:Reference number: Z19022
A:Accession: T18787
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-601 <WIL>
A:Cross-references: UNIPROT:O17552; EMBL:Z93372; PIDN:CAB07545.1; GSPDB:GN00021; CESP:BE1
A:Experimental source: clone BE10
C:Genetics:
A:Gene: CESP:BE10.3
A:Map position: 3
A:Introns: 21/3; 54/3; 117/2; 157/1; 212/2; 533/3

Query Match 82.5%; Score 33; DB 2; Length 601;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 453 NSKDYFH 459

RESULT 13
T41943
major DNA binding protein - human herpesvirus 7 (strain JI)
C:Species: human herpesvirus 7
A:Variety: strain JI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T41943
R:Nicholas, J.

```


submitted to the EMBL Data Library, December 1995

A;Description: Determination and analysis of the complete nucleotide sequence of human h
A;Reference number: Z22022
A;Accession: T41943
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1131 <NIC>
A;Cross-references: UNIPROT:P52339; EMBL:U43400; PIDN:AAC54703.1
A;Experimental source: strain J1
C;Genetics:
A;Note: U41
C;Superfamily: herpesvirus DNA-binding protein

Query Match 82.5%; Score 33; DB 2; Length 1131;

Best Local Similarity 71.4%; Pred. No. 1.1e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7

Db 943 ASRDYFH 949

RESULT 14

T00252

MEGF1 protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C;Accession: T00252

R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.

Genomics 51, 27-34, 1998

A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs

A;Reference number: Z14126; MUID:98360089; PMID:9693030

A;Accession: T00252

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-4351 <NAK>

A;Cross-references: UNIPROT:O88277; EMBL:AB011527; NID:G3449285; PIDN:BA32458.1; PID:G3

A;Experimental source: brain

C;Genetics:

A;Gene: MEGF1

C;Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G repe

F;3798-3949/Domain: laminin G repeat homology <LGR>

F;3953-3985/Domain: EGF homology <EGF>

F;3992-4023/Domain: EGF homology <EGF1>

Query Match 82.5%; Score 33; DB 2; Length 4351;

Best Local Similarity 71.4%; Pred. No. 4.6e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYFH 7

Db 605 NEQDYFH 611

RESULT 15

JC7121

androgenic gland hormone precursor - common pill bug

C;Species: Armadillidium vulgare (common pill bug)

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C;Accession: JC7121

R;Okuno, A.; Hasegawa, Y.; Ohira, T.; Katakura, Y.; Nagasawa, H.

Biochem. Biophys. Res. Commun. 264, 419-423, 1999

A;Title: Characterization and cDNA cloning of androgenic gland hormone of the terrestria

A;Reference number: JC7121; MUID:20001935; PMID:10529379

A;Accession: JC7121

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-144 <OKU>

A;Cross-references: UNIPROT:Q9U8R2; DBJ:AB029615; GB:AB029615; NID:G6446571; PID:G64465

A;Experimental source: androgenic gland

Query Match

Best Local Similarity 80.0%; Score 32; DB 2; Length 144;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY

3 QDYFH 7

Db

82 QDYFH 86

Search completed: February 8, 2005, 18:43:03

Job time : 2.81081 secs

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:25:28 ; Search time 4.10811 Seconds
(without alignments)
872.556 Million cell updates/sec

Title: US-10-820-200-2_COPY_161_167
Perfect score: 40
Sequence: 1 SSQDPFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	498	1 AMYA ASPAW	Q02905 aspergillus
2	40	100.0	498	2 Q76CT3	Q76ct3 aspergillus
3	40	100.0	499	1 AMYA ASPOR	P10529 aspergillus
4	40	100.0	499	1 AMYB ASPAW	Q02906 aspergillus
5	40	100.0	499	1 AMY ASPSH	P30292 aspergillus
6	40	100.0	499	2 Q6FTH4	Q96th4 aspergillus
7	40	100.0	499	2 Q76L99	Q76l99 aspergillus
8	40	100.0	499	2 Q7LV45	Q7lv45 aspergillus
9	40	100.0	624	1 AMY1 LIPKO	Q01117 lipomyces k
10	40	100.0	647	2 Q6YF33	Q6yF33 lipomyces s
11	37	92.5	224	2 Q70BE3	Q70be3 pseudomonas
12	36	90.0	596	2 Q7RL16	Q7rl16 plasmodium
13	36	90.0	627	2 Q6FNF7	Q6fnf7 candida gla
14	35	87.5	760	2 Q7R038	Q7r038 giardia lam
15	35	87.5	760	2 Q9XZJ0	Q9xzj0 giardia lam
16	34	85.0	130	2 Q23373	Q23373 arabidopsis
17	34	85.0	7048	2 Q6D739	Q6d739 erwinia car
18	33	82.5	125	2 Q75LX1	Q75lx1 ashbya goss
19	33	82.5	193	2 Q8LMS1	Q8lms1 oryza sativ
20	33	82.5	248	2 Q7YSZ6	Q7ysz6 rhodnius pr
21	33	82.5	257	2 Q6P396	Q6p396 mus musculu
22	33	82.5	370	2 Q8BYC0	Q8byc0 mus musculu
23	33	82.5	420	2 Q39309	Q39309 equid herpe
24	33	82.5	420	2 Q55525	Q55525 equid herpe
25	33	82.5	585	1 CTC9 MOUSE	Q9d718 mus musculu
26	33	82.5	601	2 Q17552	Q17552 caenorhabdi
27	33	82.5	762	2 Q7NAH6	Q7nah6 mycoplasma
28	33	82.5	1131	1 DNBI HHV7J	P52339 human herpe
29	33	82.5	1131	2 Q56282	Q56282 human herpe
30	33	82.5	1304	2 Q6FR07	Q6fr07 candida gla
31	33	82.5	1731	2 Q9P230	Q9p230 homo sapien

32 33 82.5 3124 2 Q96L91 Q96l91 homo sapien
33 33 82.5 4351 1 FAT2 RAT Q88277 rattus norv
34 32 80.0 49 2 Q6XIZ7 Q6xy77 spiroplasma
35 32 80.0 95 2 Q7RM89 Q7rm89 plasmodium
36 32 80.0 144 1 AGH ARMVU Q9u8r2 armadillidi
37 32 80.0 163 2 Q9CUQ8 Q9cuq8 mus musculu
38 32 80.0 180 2 Q97SK2 Q97ak2 streptococc
39 32 80.0 181 2 Q7USU1 Q7ueu1 rhodopirell
40 32 80.0 183 2 Q8DZ39 Q8dz39 streptococc
41 32 80.0 183 2 Q8B4P7 Q8eb4p7 streptococc
42 32 80.0 215 2 P97975 P97975 unidentifie
43 32 80.0 218 2 Q9CUH8 Q9cuH8 mus musculu
44 32 80.0 255 2 Q8EU15 Q8eu15 oceanobacil
45 32 80.0 261 2 Q7JL56 Q73l56 treponema d

ALIGNMENTS

RESULT 1
ID AMYA ASPAW STANDARD; PRT; 498 AA.
AC Q02905;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase A).
GN Name=AMYA;
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L., Kodama K.H.,
Royer T.J., Thompson S.A., Ward M., Wilson L.J., Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52755; CAA36966.1; -.
CC PIR; A48305; A48305.
CC HSSP; P10529; 7TAA.
CC InterPro; IPR006589; Alp_amyl_cat_sub.
CC InterPro; IPR006047; Alpha_amyl_cat.
CC InterPro; IPR006046; Glyco_hydro_13.
CC Pfam; PF00128; Alpha-amylase; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Amy; 1.
KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolyase; Multigene family; Signal.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 498 Alpha-amylase A.
FT ACT_SITE 227 227 Nucleophile (By similarity).
FT ACT_SITE 251 251 Proton donor (By similarity).
FT ACT_SITE 318 318 By similarity.
FT METAL 142 142 Calcium 1 (By similarity).

```

FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
FT METAL 196 196 similarity).
FT METAL 227 227 Calcium 1 (By similarity).
FT METAL 231 231 Calcium 2 (By similarity).
FT METAL 251 251 Calcium 2 (By similarity).
FT DISULFID 51 59 By similarity.
FT DISULFID 171 185 By similarity.
FT DISULFID 261 304 By similarity.
FT DISULFID 461 496 By similarity.
FT CARBOHYD 218 218 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 498 AA; 54880 MW; 7658511BC01A8A01 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 498;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
Db 162 SSQDYFH 168

RESULT 2
Q76CT3 PRELIMINARY; PRT; 498 AA.
ID Q76CT3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amylase.
GN Name=amyA;
OS Aspergillus kawachi (Aspergillus awamori var. kawachi).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=40384;
RN [1]
RP SEQUENCE FROM N.A.
RA Ito K.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB109452; BAD01051.1; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Amyy; 1.
DR SEQUENCE 498 AA; 54679 MW; 8407BBFC7F4E3FC0 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
Db 161 SSQDYFH 167

RESULT 3
AMYA_ASPOR STANDARD; PRT; 499 AA.
AC P10529; P11763; Q00250;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 25-OCT-2004 (Rel. 12, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA) (1,4-
DE Alpha-D-glucan glucanohydrolase).
GN Name=AMY1;
GN and
GN Name=AMY2;

```

```

GN Name=AMY3;
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 63303;
RX MEDLINE=89237897; PubMed=2785629;
RA Wirsel S., Lachmund A., Wildhardt G., Rutkowski E.;
RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
RT intron-exon organization.";
RL Mol. Microbiol. 3:3-14(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89378767; PubMed=2789162; DOI=10.1016/0378-1119(89)90096-6;
RA Genes M.J., Dove M.J., Sellig V.L.;
RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
RT containing eight introns.";
RL Gene 79:107-117(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene of
RT Aspergillus oryzae.";
RL Agric. Biol. Chem. 53:593-599(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128276; PubMed=2612911; DOI=10.1016/0378-1119(89)90506-4;
RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A., Uda S.;
RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
RT evidence for multiple related genes.";
RL Gene 84:319-327(1989).
RN [5]
RP SEQUENCE OF 22-499.
RA Toda H., Kondo K., Narita K.;
RT "The complete amino acid sequence of Taka-amylase A.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
RN [6]
RP SEQUENCE OF 206-225.
RX MEDLINE=74001521; PubMed=4733850;
RA Isemura S., Ikenaka T.;
RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
RT A with trypsin and chymotrypsin.";
RL J. Biochem. 74:1-10(1973).
RN [7]
RP SEQUENCE OF 433-499.
RA Narita K.;
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=80227691; PubMed=6156152;
RA Matsuura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
RA Toda H., Narita K., Kakudo M.;
RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
RT 3-A resolution.";
RL J. Biochem. 87:1555-1558(1980).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=84212370; PubMed=6609921;
RA Matsuura Y., Kusunoki M., Harada W., Kakudo M.;
RT "Structure and possible catalytic residues of Taka-amylase A.";
RL J. Biochem. 95:697-702(1984).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
RX MEDLINE=97428212; PubMed=9283074; DOI=10.1021/bi970539i;
RA Brzozowski A.M., Davies G.J.;
RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
RT inhibitor acarbose at 2.0-A resolution.";
RL Biochemistry 36:10837-10845(1997).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.

```

CC -1- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
 CC at high concentrations.
 CC -1- SUBUNIT: Monomer.
 CC -1- BIOTECHNOLOGY: Used in the brewing industry to increase the
 CC fermentability of beer worts (including those made from unmalted
 CC cereals), in the starch industry to make high maltose and high DE
 CC syrups (starch saccharification), in the alcohol industry to
 CC reduce fermentation time, in the cereal food industry for flour
 CC supplementation and improvement of chilled and frozen dough, and
 CC in the forestry industry for low-temperature modification of
 CC starch. Sold under the name Fungamyl by Novozymes.
 CC -1- MISCELLANEOUS: The sequence of Amy1 and Amy2 is shown.
 CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X12725; CAA31218.1; --
 CC EMBL; X12726; CAA31219.1; --
 CC EMBL; X12727; CAA31220.1; --
 CC EMBL; D00434; BAA00336.1; --
 CC EMBL; M33218; AAA32708.1; --
 CC PIR; JK0201; JK0201.
 CC PIR; JT0466; JT0466.
 CC PIR; S04548; ALAS1.
 CC PDB; 2TAA; X-ray; A=22-499.
 CC PDB; 6TAA; X-ray; @=22-499.
 CC PDB; 7TAA; X-ray; @=22-499.
 CC GlycoSuiteDB; P10529; --
 CC InterPro; IPR006589; Alp_amy1_cat_sub.
 CC InterPro; IPR006047; Alpha_amy1_cat.
 CC Pfam; PF00128; Alpha-amy1ase; 1.
 CC PRINTS; PR00110; ALPHAAMYLASE.
 CC SMART; SM00642; Amy; 1.
 CC 3D-structure; Calcium-binding; Carbohydrate metabolism;
 CC Direct protein sequencing; Glycoprotein; Glycosidase; Hydrolase;
 CC Multigene family; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 499 Alpha-amy1ase A.
 FT ACT_SITE 227 227 Nucleophile.
 FT ACT_SITE 251 251 Proton donor.
 FT ACT_SITE 318 318
 FT METAL 142 142 Calcium 1.
 FT METAL 183 183 Calcium 1 (via carbonyl oxygen).
 FT METAL 196 196 Calcium 1.
 FT METAL 227 227 Calcium 2.
 FT METAL 231 231 Calcium 1 (via carbonyl oxygen).
 FT METAL 251 251 Calcium 2.
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218 N-linked (GlcNAc...).
 FT VARIANT 56 56 /FTID=CAR_000125.
 FT VARIANT 172 172 Q -> R (in Amy3).
 FT CONFLICT 93 94 F -> L (in Amy3).
 FT CONFLICT 106 106 TT -> DC (in Ref. 5).
 FT CONFLICT 184 184 D -> Y (in Ref. 5).
 FT CONFLICT 195 195 P -> L (in Ref. 3).
 FT CONFLICT 255 255 G -> V (in Ref. 3).
 FT CONFLICT 291 291 D -> H (in Ref. 4).
 FT CONFLICT 345 345 I -> L (in Ref. 5).
 FT CONFLICT 370 370 L -> A (in Ref. 4).
 FT CONFLICT 406 409 WPIV -> PYI (in Ref. 5).
 FT CONFLICT 448 448 G -> S (in Ref. 5).
 FT CONFLICT 497 497 S -> SD (in Ref. 5 and 7).

FT HELIX 24 27
 FT TURN 28 29
 FT STRAND 32 35
 FT HELIX 37 40
 FT STRAND 42 42
 FT TURN 43 44
 FT HELIX 53 55
 FT STRAND 61 61
 FT HELIX 63 68
 FT TURN 69 69
 FT HELIX 70 74
 FT TURN 75 77
 FT STRAND 80 83
 FT STRAND 87 89
 FT STRAND 94 94
 FT TURN 95 96
 FT STRAND 97 97
 FT TURN 100 101
 FT STRAND 105 111
 FT TURN 113 114
 FT HELIX 118 130
 FT TURN 131 132
 FT STRAND 134 139
 FT STRAND 143 143
 FT STRAND 146 148
 FT HELIX 150 152
 FT HELIX 155 157
 FT STRAND 159 159
 FT HELIX 164 166
 FT STRAND 167 167
 FT STRAND 172 172
 FT TURN 176 177
 FT HELIX 179 184
 FT STRAND 186 188
 FT STRAND 192 194
 FT TURN 199 200
 FT TURN 202 219
 FT TURN 220 220
 FT STRAND 223 226
 FT HELIX 229 231
 FT TURN 234 235
 FT HELIX 236 244
 FT TURN 245 245

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
 Db 162 SSQDYFH 168

RESULT 4
 AMYB ASPAW STANDARD; PRT; 499 AA.
 ID AMYB ASPAW
 AC Q02906;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alpha-amy1ase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase B).
 GN Name=AMYB;
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UVK143F;
 RX MEDLINE=90254827; PubMed=2340591;
 RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L., Kodama K.H.,

RA Royer T.J., Thompson S.A., Ward M., Wilson L.J., Berka R.M.;
 RT "Cloning, characterization, and expression of two alpha-amylase genes
 RL from *Aspergillus niger* var. *awamori*.";
 CC Curr. Genet. 17:203-212(1990).
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
 CC at high concentrations (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52756; CAA36967.1; -;
 DR PIR; B48305; B48305.
 DR HSSP; P10529; 7TAA.
 DR InterPro; IPR006589; Alp_amyl_cat_sub.
 DR InterPro; IPR006047; Alpha_amyl_cat.
 DR InterPro; IPR006046; Glyco_hydro_13.
 DR Pfam; PF00128; Alpha-amylase; 1.
 DR PRINTS; PR00110; ALPHAAMYLASE.
 DR SMART; SM00642; Aamy; 1.
 KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
 KW Hydrolase; Multigene family; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 499 Alpha-amylase B.
 FT ACT_SITE 227 227 Nucleophile (By similarity).
 FT ACT_SITE 251 251 Proton donor (By similarity).
 FT ACT_SITE 318 318 By similarity.
 FT ACT_SITE 318 318 By similarity.
 FT METAL 142 142 Calcium 1 (By similarity).
 FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
 FT similarity).
 FT METAL 196 196 Calcium 1 (By similarity).
 FT METAL 227 227 Calcium 2 (By similarity).
 FT METAL 231 231 Calcium 1 (via carbonyl oxygen) (By
 FT similarity).
 FT METAL 251 251 Calcium 2 (By similarity).
 FT DISULFID 51 59 By similarity.
 FT DISULFID 171 185 By similarity.
 FT DISULFID 261 304 By similarity.
 FT DISULFID 461 496 By similarity.
 FT CARBOHYD 218 218 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;
 Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSQDYFH 7
 Db 162 SSQDYFH 168
 RESULT 5
 ID AMY ASPSH STANDARD; PRT; 499 AA.
 AC P30292;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase).
 DE Name=AMY;
 GN *Aspergillus shirousami*.
 OS *Aspergillus shirousami*.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
 OX NCBI_TaxID=5070;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE=92323146; PubMed=1368777;
 RX Shibuya I., Tamura G., Ishikawa T., Hara S.;
 RT "Cloning of the alpha-amylase cDNA of *Aspergillus shirousami* and its
 RL expression in *Saccharomyces cerevisiae*.";
 RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
 CC at high concentrations (By similarity).
 CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D10461; BRA01255.1; -;
 DR HSSP; P10529; 7TAA.
 DR InterPro; IPR006589; Alp_amyl_cat_sub.
 DR InterPro; IPR006047; Alpha_amyl_cat.
 DR InterPro; IPR006046; Glyco_hydro_13.
 DR Pfam; PF00128; Alpha-amylase; 1.
 DR PRINTS; PR00110; ALPHAAMYLASE.
 DR SMART; SM00642; Aamy; 1.
 KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
 KW Hydrolase; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 499 Alpha-amylase.
 FT ACT_SITE 227 227 Nucleophile (By similarity).
 FT ACT_SITE 251 251 Proton donor (By similarity).
 FT ACT_SITE 318 318 By similarity.
 FT ACT_SITE 318 318 By similarity.
 FT METAL 142 142 Calcium 1 (By similarity).
 FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
 FT similarity).
 FT METAL 196 196 Calcium 1 (By similarity).
 FT METAL 227 227 Calcium 2 (By similarity).
 FT METAL 231 231 Calcium 1 (via carbonyl oxygen) (By
 FT similarity).
 FT METAL 251 251 Calcium 2 (By similarity).
 FT DISULFID 51 59 By similarity.
 FT DISULFID 171 185 By similarity.
 FT DISULFID 261 304 By similarity.
 FT DISULFID 461 496 By similarity.
 FT CARBOHYD 218 218 N-linked (GlcNAc...) (By similarity).
 SQ SEQUENCE 499 AA; 54852 MW; 1FB7AB50DA01C03F CRC64;
 Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSQDYFH 7
 Db 162 SSQDYFH 168
 RESULT 6
 ID Q96TH4 PRELIMINARY; PRT; 499 AA.
 AC Q96TH4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Taka-amylase A (EC 3.2.1.1).
 DE Name=amyA;
 GN *Aspergillus oryzae*.
 OS *Aspergillus oryzae*.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
 OX NCBI_TaxID=5062;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=RIB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amyolytic gene expression in Aspergillus
RT oryzae.";
RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB021876; BAA95703.1; -.
DR PIR; JN0588; JN0588.
DR PIR; S04549; ALAS3.
DR HSP; P10529; TPA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Aamy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 162 SSQDYFH 168

RESULT 7
Q76L99 PRELIMINARY; PRT; 499 AA.
AC Q76L99;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amylase precursor.
GN Name=amyI;
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT-11;
RA Matsubara T.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB083159; BAD06002.1; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Aamy; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 499 AA; 54794 MW; 2B357AE38B36C1C2 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 162 SSQDYFH 168

RESULT 8
Q7LV45 PRELIMINARY; PRT; 499 AA.
AC Q7LV45;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amylase (EC 3.2.1.1).
GN Name=amyI;
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5059;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86-10D;
RA Fakhoury A.M., Woloshuk C.P.;
RT "AmyI, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914(1999).
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AF13925; AAF14264.1; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Aamy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 162 SSQDYFH 168

RESULT 9
AMYI_LIPKO STANDARD; PRT; 624 AA.
AC Q01117;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE Glucanohydrolase 1).
GN Name=LkAl;
OS Lipomyces kononenkoae.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Lipomycetaceae; Lipomyces.
OX NCBI_TaxID=34357;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IGC4052B;
RX MEDLINE=96105202; PubMed=8529895; DOI=10.1016/0378-1119(95)00633-0;
RA Steyn A.J.C., Marmur J., Pretorius I.S.;
RT "Cloning, sequence analysis and expression in yeasts of a cDNA
RT containing a Lipomyces kononenkoae alpha-amylase-encoding gene.";
RL Gene 166:65-71(1995).
RN [2]
RP SEQUENCE OF 29-44.
RC STRAIN=IGC4052B;
RX MEDLINE=96132108; PubMed=8593683;
RA Steyn A.J.C., Pretorius I.S.;

```

RT "Characterization of a novel alpha-amylase from *Lipomyces kononenkoae*
RL and expression of its gene (UKA1) in *Saccharomyces cerevisiae*."; Curr. Genet. 28:526-533 (1995).
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -|- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC -|- SIMILARITY: Contains 1 carbohydrate binding type-21 (CBM21)
CC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: U30376; AAC49622.1; ALT_INIT.
DR PIR: JC4510; JC4510.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR006589; Alp_amy1_cat_sub.
DR InterPro: IPR006047; Alpha_amy1_cat.
DR InterPro: IPR005036; CBM 21.
DR Pfam: PF00128; Alpha-amyLase; 1.
DR Pfam: PF03370; CBM 21; 1.
DR SMART: SM00642; Amy; 1.
DR KX Calcium-binding; Carbohydrate metabolism; Direct protein sequencing;
KW Glycoprotein; Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 28
FT CHAIN 29 624 Alpha-amyLase 1.
FT DOMAIN 40 133 Carbohydrate binding type-21.
FT ACT_SITE 353 353 Nucleophile (By similarity).
FT ACT_SITE 377 377 Proton donor (By similarity).
FT ACT_SITE 444 444 By similarity.
FT METAL 268 268 Calcium 1 (By similarity).
FT METAL 309 309 Calcium 1 (via carbonyl oxygen) (By
FT similarity).
FT METAL 322 322 Calcium 1 (By similarity).
FT METAL 353 353 Calcium 2 (By similarity).
FT METAL 357 357 Calcium 1 (via carbonyl oxygen) (By
FT similarity).
FT METAL 377 377 Calcium 2 (By similarity).
FT DISULFID 177 185 By similarity.
FT DISULFID 297 311 By similarity.
FT DISULFID 387 430 By similarity.
FT DISULFID 587 622 By similarity.
FT CARBOHYD 304 304 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 344 344 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 624 AA; 68876 MW; 87EB16534F5A9A9F CRC64;

Query Match 100.0%; Score 40; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
Db 288 SSQDYFH 294
|:|||||

RESULT 10
ID Q6YF33 PRELIMINARY; PRT; 647 AA.
AC Q6YF33;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amyLase
OS *Lipomyces starkeyi*.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Lipomycetaceae; Lipomyces.

OX NCBI_TaxID=29829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSM 22M;
RX PubMed=15043869; DOI=10.1016/j.femsle.2004.01.036;
RA Kang H.K., Lee J.H., Kim D., Day D.F., Robyt J.F., Park K.-H.,
RA Moon T.-W.;
RT "Cloning and expression of *Lipomyces starkeyi* alpha-amyLase in
RT *Escherichia coli* and determination of some of its properties."; FEMS Microbiol. Lett. 233:53-64 (2004).
RL EMBL: AY155463; AAN75021.1; -.
DR HSSP: P26827; 1A47.
DR GO: GO:0004556; F:alpha-amyLase activity; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR006047; Alpha_amy1_cat.
DR InterPro: IPR006589; Alp_amy1_cat_sub.
DR InterPro: IPR005036; CBM 21.
DR Pfam: PF00128; Alpha-amyLase; 1.
DR Pfam: PF03370; CBM 21; 1.
DR SMART: SM00642; Amy; 1.
SQ SEQUENCE 647 AA; 70957 MW; D946071DB7560144 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 647;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
Db 311 SSQDYFH 317
|:|||||

RESULT 11
ID Q70BE3 PRELIMINARY; PRT; 224 AA.
AC Q70BE3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative helicase.
OS *Pseudomonas syringae* (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RA Panelli V., Finetti-Sialer M., Gallitelli D.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ604569; CAE34310.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00271; Helicase_C; 1.
KW Helicase.
SQ SEQUENCE 224 AA; 25431 MW; 8701280423194015 CRC64;

Query Match 92.5%; Score 37; DB 2; Length 224;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
Db 130 STQDYFH 136
|:|||||

RESULT 12
ID Q7RL16 PRELIMINARY; PRT; 596 AA.
AC Q7RL16;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE *Drosophila melanogaster* CG11926 gene product, putative

DE (Fragment).
GN Names=PY02732;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XN;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angluoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000756; EAA22214.1; -.
DR InterPro; IPR004353; Yeast73DUF.
DR Pfam; PF03164; DUF254; 1.
DR PRINTS; PR01546; YEAST73DUF.
FT NON TER 1
SQ SEQUENCE 596 AA; 65216 MW; 8170C2D33F53117D CRC64;

Query Match 90.0%; Score 36; DB 2; Length 596;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 446 SSQDPFH 452
|||||

RESULT 13
Q6FN7 PRELIMINARY; PRT; 627 AA.
AC Q6FN7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Candida glabrata strain CBS138 chromosome K complete sequence.
GN ORFNames=CAGL0K002979;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Caaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anhouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri A.,
RA Kereest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolaiki M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pallenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Senenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."

RL Nature 430:35-44(2004).
DR EMBL; CR380957; CAG61198.1; -.
DR InterPro; IPR002013; SyJa_N.
DR Pfam; PF02383; SyJa_N; 1.
DR PROSITE; PS0275; SAC; 1.
SQ SEQUENCE 627 AA; 71471 MW; DEAOB02466F4F480 CRC64;

Query Match 90.0%; Score 36; DB 2; Length 627;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 365 SSQDPFH 371
|||||

RESULT 14
Q7R038 PRELIMINARY; PRT; 760 AA.
AC Q7R038;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 456 15756 18038.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000038; EAA40649.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR PRINTS; PR01415; ANKYRIN.
SQ SEQUENCE 760 AA; 85474 MW; A01440ACE754099C CRC64;

Query Match 87.5%; Score 35; DB 2; Length 760;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 5 SSQDPFH 11
|||||

RESULT 15
Q9XZJ0 PRELIMINARY; PRT; 760 AA.
AC Q9XZJ0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ankyrin-like protein.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Portland;
RA Sekyere E.O., Edwards M.R., Stewart T.S., Schofield P.J., Knodler L.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF144322; AAD28486.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.

KW ANK repeat.
SQ SEQUENCE 760 AA; 85542 MW; 12E54841019CD475 CRC64;
Query Match 87.5%; Score 35; DB 2; Length 760;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSQDYFH 7
| | | | |
Db 5 SSQDWFH 11

Search completed: February 8, 2005, 18:41:55
Job time : 6.10811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:22:38 ; Search time 7.45367 Seconds
(without alignments)
674.552 Million cell updates/sec

Title: US-10-820-200-2_COPY_98_110

Perfect score: 78

Sequence: 1 AHGYWQDIYSL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	478	2	AAR46065 Mutant al
2	78	100.0	478	2	AAR72450 Aspergill
3	78	100.0	478	2	AAR78270 Aspergill
4	78	100.0	478	2	AAW14500 Aspergill
5	78	100.0	498	4	AAW14500 Amino aci
6	78	100.0	498	4	AAW14500 Amino aci
7	78	100.0	498	8	ADT89632 Aspergill
8	78	100.0	498	8	ADT89632 Aspergill
9	78	100.0	1095	6	ABP96630 Alpha-amy
10	75	96.2	495	6	ABB80177 A. fumiga
11	70	89.7	423	4	ABB09072 Aspergill
12	69	88.5	494	6	ABP97894 Amino aci
13	69	88.5	493	2	AAR88213 Alpha-amy
14	69	88.5	500	8	ADN19570 Bacterial
15	67	85.9	513	8	ADN19570 Bacterial
16	65	83.3	547	6	ABB80179 A. fumiga
17	64	82.1	484	5	AAE24207 Aspergill
18	64	82.1	484	8	ADT89632 Aspergill
19	63	80.8	555	6	ABP97896 Amino aci
20	63	80.8	567	6	ABP97899 Amino aci
21	62	79.5	478	2	AAW14500 Amino aci
22	58	74.4	468	2	AAW14500 Amino aci
23	58	74.4	494	1	AAW14500 Amino aci
24	58	74.4	630	6	ABB80178 A. fumiga
25	57	73.1	549	6	ABP97895 Amino aci

26	56	71.8	493	8	ADS44293	Ade44293 Bacterial
27	54	69.2	468	2	AAR63184	Aar63184 Variant a
28	53	67.9	468	2	AAR63185	Aar63185 Variant a
29	53	67.9	564	8	ADS44249	Ade44249 Bacterial
30	52	66.7	711	5	ABP26024	Abp26024 Streptoco
31	50	64.1	452	2	AAR63186	Aar63186 Variant a
32	50	64.1	511	1	AAW17605	AAW17605 Bacillus
33	50	64.1	511	1	AAW17605	AAW17605 Bacillus
34	50	64.1	512	2	AAW17605	AAW17605 Bacillus
35	49	62.8	468	2	AAR63187	Aar63187 Variant a
36	49	62.8	556	2	AAW17599	AAW17599 Bacillus
37	49	62.8	556	2	AAW17599	AAW17599 Bacillus
38	49	62.8	556	2	AAW17599	AAW17599 Bacillus
39	49	62.8	556	2	AAW17599	AAW17599 Bacillus
40	49	62.8	556	2	AAW17599	AAW17599 Bacillus
41	49	62.8	556	2	AAW17599	AAW17599 Bacillus
42	49	62.8	556	2	AAW17599	AAW17599 Bacillus
43	49	62.8	556	2	AAW17599	AAW17599 Bacillus
44	49	62.8	556	2	AAW17599	AAW17599 Bacillus
45	49	62.8	556	2	AAW17599	AAW17599 Bacillus

ALIGNMENTS

RESULT 1

AAR46065
ID AAR46065 standard; protein; 478 AA.

XX AAR46065;

XX AC

XX 25-MAR-2003 (revised)

DT 18-JUL-1994 (first entry)

XX XX

DE Mutant alpha-amylase.

XX XX

KW Methionine substitution; stability; activity; detergent;

KW dishwashing agents; liquefaction agents.

XX XX

OS Aspergillus oryzae.

XX XX

PN WO9402597-A1.

XX XX

PD 03-FEB-1994.

XX XX

PF 06-JUL-1993; 93WO-DK000230.

XX XX

PR 23-JUL-1992; 92DK-00000946.

PR 16-DEC-1992; 92DK-00001503.

PR 15-MAR-1993; 93DK-00000292.

XX XX

PA (NOVO) NOVO-NORDISK AS.

XX XX

PI Svendsen A, Bisgard-Frantzen H;

XX XX

DR WPI; 1994-048855/06.

XX XX

PT Mutant alpha-amylase from Bacillus species comprising a methionine substitution - with improved stability and activity at low pH, for use in detergents, dishwashing agents and liquefaction agents.

XX XX

PS Claim 1; Page 7; 20pp; English.

XX XX

CC The sequence as that of the Aspergillus oryzae alpha amylase, sold commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can be mutated by substitution of one or more of its methionine residues for any amino acid other than cysteine. The mutant alpha-amylase exhibits a better activity level and better stability in the presence of oxidising agents than previous mutant alpha amylases, and improved thermostability at moderately low pH. The enzyme can be used as an additive for detergents, dishwashing agents and liquefaction agents. (Updated on 25-MAR-2003 to correct FN field.)

XX

RESULT 3
AAR78270
ID AAR

FH Key Location/Qualifiers
 FT Misc-difference 13..45
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 7-23 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 33"
 FT
 FT Misc-difference 14..40
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 8-18 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to this
 FT fragment; claim 35"
 FT
 FT Misc-difference 28..42
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 12-19 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 30"
 FT
 FT Misc-difference 32..38
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 14-15 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to this
 FT fragment; claim 32"
 FT
 FT Misc-difference 66..84
 FT /label= loop 2 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 44-57 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 18"
 FT
 FT Misc-difference 70..78
 FT /label= loop 2 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 48-51 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to this
 FT fragment; claim 20"
 FT
 FT Misc-difference 98..210
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 117-185 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 24"
 FT
 FT Misc-difference 102..206
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 121-181 of AAW14499 is
 FT deleted or replaced with a fragment corresponding to this
 FT fragment; claim 26"
 FT
 FT Misc-difference 121..181
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to this fragment is deleted or replaced with a fragment
 FT corresponding to 102-206 of AAW14499; claim 41"
 FT
 FT Misc-difference 121..174
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to this fragment is deleted or
 FT replaced with a fragment corresponding to 102-199 of
 FT AAW14499; claim 42"
 FT
 FT Misc-difference 165..177
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 195-202 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 21"
 FT
 FT Misc-difference 166..173
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 196-198 of AAW14499 is
 FT deleted or replaced with a fragment corresponding to this
 FT fragment; claim 23"
 FT
 FT Misc-difference 181..184
 FT /note= "an amino acid fragment corresponding to this
 FT region is deleted from the parent sequence of a variant
 FT Fungamyl; claim 43"
 FT
 FT Misc-difference 291..313
 FT /label= loop 8 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 322-346 of AAW14498 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 36"
 FT
 FT Misc-difference 297..313
 FT /label= loop 8 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 325-345 of AAW14498 is
 FT deleted or replaced with a fragment corresponding to this
 FT fragment; claim 38"
 FT
 FT XX WO9623874-A1.
 FT PN
 FT XX
 FT PD 08-AUG-1996.
 FT XX
 FT XX 05-FEB-1996; 96WO-DK0000057.
 FT PF
 FT XX 03-FEB-1995; 95DK-00000128.
 FT PR
 FT PR 23-OCT-1995; 95DK-00001192.
 FT PR 10-NOV-1995; 95DK-00001256.
 FT XX
 FT PA (NOVO) NOVO-NORDISK AS.
 FT XX
 FT PI Svendsen A, Bisgard-Frantzen H, Borchert TV;
 FT XX
 FT DR WPI; 1996-371424/37.
 FT XX
 FT PT Alpha-amylase variants and methods of production - have altered
 FT properties such as calcium dependency, substrate binding and stability.
 FT
 FT XX Disclosure; Page 87-88; 171pp; English.
 FT PS
 FT XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
 FT CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
 FT CC methods of constructing them) are claimed. Examples of variants are
 FT CC featured above. The variants have altered properties such as calcium
 FT CC dependency, substrate binding and stability. Also one or more proline or
 FT CC cysteine residues in the variant is modified or replaced with a non-
 FT CC proline or non-cysteine residue such as alanine. The variants can be used
 FT CC for (dish)washing, as detergent additives or for fabric desizing or
 FT CC starch liquefaction. They can also be used for the production of
 FT CC sweeteners and ethanol from starch. See also AAW14498-99
 FT XX
 FT SQ Sequence 478 AA;
 FT
 FT Query Match 100.0%; Score 78; DB 2; Length 478;
 FT Best Local Similarity 100.0%; Pred. No. 0.0011;
 FT Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 FT
 FT QY 1 AYHGWQDDIYSL 13
 FT |||||
 FT 78 AYHGWQDDIYSL 90
 FT Db
 FT
 FT RESULT 5
 FT AAB84206
 FT ID AAB84206 standard; protein; 498 AA.
 FT XX
 FT AC AAB84206;
 FT XX
 FT DT 06-AUG-2001 (first entry)
 FT XX

DE Amino acid sequence of a fungamyl-like alpha-amylase.
 XX Fungamyl-like alpha-amylase; Glucoamylase; Dextrinisation; maltose;
 KW alcohol; starch; dough improver; brewing; starch liquification.
 XX Aspergillus oryzae.
 OS
 PN WO200134784-A1.
 XX
 PD 17-MAY-2001.
 XX
 PF 10-NOV-2000; 2000WO-DK000626.
 XX
 PR 10-NOV-1999; 99DK-00001617.
 XX
 PA (NOVO) NOVOZYMES AS.
 XX
 XX Bisgard-Frantzen H, Svendsen A, Pedersen S;
 PI
 XX WPI; 2001-367478/38.
 DR N-PSDB; AAF90208.
 XX
 XX New variant of Fungamyl-like alpha-amylase, useful for production of
 PT maltose syrups, includes mutations that improve stability against heat
 PT and acidic pH.
 XX
 PS Claim 1; Page 42-45; 49pp; English.
 XX
 CC The present sequence represents a fungamyl-like alpha-amylase. The
 CC specification describes variants of this fungamyl-like alpha-amylase,
 CC which have an alteration in one the amino acid regions 98-110, 150-160,
 CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion or
 CC substitution of an amino acid or an insertion of an amino acid downstream
 CC of a particular position. The variants retain alpha-amylase activity, and
 CC have better heat stability and/or stability at acidic pH, relative to
 CC wild-type enzyme. The variants can therefore be used at higher
 CC temperatures (more efficient conversion or faster reaction, and have
 CC reduced need for cooling and reduced risk of contamination). The variants
 CC may also be used in conjunction with other enzymes, particularly
 CC Glucoamylase during dextrinisation. The variants are used to produce
 CC syrups, particularly of high maltose content, or alcohol, from starch; as
 CC dough improver for baked goods; in brewing, to increase fermentability of
 CC the wort; and for liquefaction of starch
 XX
 SQ Sequence 498 AA;
 Query Match 100.0%; Score 78; DB 4; Length 498;
 Best Local Similarity 100.0%; Pred. NO. 0.0011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AYHGYWQDDIYSL 13
 DB 98 AYHGYWQDDIYSL 110
 RESULT 6
 ADT89632
 ID ADT89632 standard; protein; 498 AA.
 XX
 AC ADT89632;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Aspergillus niger neutral alpha-amylase A (amyA) protein.
 XX
 KW Glucoamylase; glaA; amyA; alpha-amylase A; enzyme.
 XX
 OS Aspergillus niger.
 XX
 PN US2004191864-A1.
 XX
 PD 30-SEP-2004.
 XX
 PF 31-MAR-2004; 2004US-00815495.
 XX
 PR 31-MAR-2003; 2003US-0459902P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 XX
 PI Connelly M, Brody H;
 XX
 XX WPI; 2004-708545/69.
 DR N-PSDB; ADT89627.
 XX
 PT Producing heterologous biological substance comprises culturing mutant of
 PT wild-type Aspergillus niger strain in medium suitable for producing
 PT heterologous biological substance and recovering heterologous biological
 PT substance.
 XX
 PS Example 10; SEQ ID NO 18; 58pp; English.

PF 31-MAR-2004; 2004US-00815495.
 XX
 PR 31-MAR-2003; 2003US-0459902P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 XX
 PI Connelly M, Brody H;
 XX
 XX WPI; 2004-708545/69.
 DR N-PSDB; ADT89631.
 XX
 PT Producing heterologous biological substance comprises culturing mutant of
 PT wild-type Aspergillus niger strain in medium suitable for producing
 PT heterologous biological substance and recovering heterologous biological
 PT substance.
 XX
 PS Example 11; SEQ ID NO 22; 58pp; English.
 XX
 CC The present invention relates to a method of producing heterologous
 CC biological substance. The method involves culturing mutant of wild-type
 CC Aspergillus niger strain in medium suitable for producing heterologous
 CC biological substance, where mutant strain comprises first nucleotide
 CC sequence encoding heterologous biological substance and second nucleotide
 CC sequences comprising modification of glucoamylase (glaA) and recovering
 CC heterologous biological substance. The present sequence is the
 CC Aspergillus niger neutral alpha-amylase A (amyA) protein.
 XX
 SQ Sequence 498 AA;
 Query Match 100.0%; Score 78; DB 8; Length 498;
 Best Local Similarity 100.0%; Pred. NO. 0.0011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AYHGYWQDDIYSL 13
 DB 99 AYHGYWQDDIYSL 111
 RESULT 7
 ADT89628
 ID ADT89628 standard; protein; 499 AA.
 XX
 AC ADT89628;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Aspergillus niger neutral alpha-amylase B (amyB) protein.
 XX
 KW Glucoamylase; glaA; amyB; alpha-amylase B; enzyme.
 XX
 OS Aspergillus niger.
 XX
 PN US2004191864-A1.
 XX
 PD 30-SEP-2004.
 XX
 PF 31-MAR-2004; 2004US-00815495.
 XX
 PR 31-MAR-2003; 2003US-0459902P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 XX
 PI Connelly M, Brody H;
 XX
 XX WPI; 2004-708545/69.
 DR N-PSDB; ADT89627.
 XX
 PT Producing heterologous biological substance comprises culturing mutant of
 PT wild-type Aspergillus niger strain in medium suitable for producing
 PT heterologous biological substance and recovering heterologous biological
 PT substance.
 XX
 PS Example 10; SEQ ID NO 18; 58pp; English.

XX The present invention relates to a method of producing heterologous
 CC biological substance. The method involves culturing mutant of wild-type
 CC *Aspergillus niger* strain in medium suitable for producing heterologous
 CC biological substance, where mutant strain comprises first nucleotide
 CC sequence encoding heterologous biological substance and second nucleotide
 CC sequences comprising modification of glucoamylase (glaA) and recovering
 CC heterologous biological substance. The present sequence is the
 CC *Aspergillus niger* neutral alpha-amylase B (amyB) protein.
 XX
 SQ Sequence 499 AA;

Query Match 100.0%; Score 78; DB 8; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
 |||||
 Db 99 AYHGYWQDDIYSL 111
 |||||

RESULT 8
 ABP96630
 ID ABP96630 standard; protein; 1095 AA.
 XX
 AC ABP96630;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Alpha-amylase/glucoamylase fusion protein sequence SEQ ID NO:45.
 XX
 KW Self-processing plant; plant; processing enzyme; alpha-amylase; grain;
 KW pullulanase; alpha-glucosidase; glucose isomerase; glucoamylase;
 KW mesophilic; thermophilic; hyperthermophilic; transgenic plant; starch;
 KW maltodextrin; ethanol; fermentation; beverage; enzyme.
 XX
 OS *Aspergillus shiroueami*.
 OS Synthetic.
 XX
 PN WO2003018766-A2.
 XX
 PD 06-MAR-2003.
 XX
 PF 27-AUG-2002; 2002WO-US027129.
 XX
 PR 27-AUG-2001; 2001US-0315281P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Lanahan MB, Basu SS, Batie CJ, Chen W, Craig J, Kinkema M;
 XX
 DR WPI; 2003-268420/26.
 DR N-PSDB; ACC44572.
 XX
 PT Novel polynucleotide encoding hyperthermophilic processing enzymes e.g.
 PT alpha-amylase, useful for producing plant to produce food products having
 PT improved taste or fermentable substrates for ethanol.
 XX
 PS Claim 1; Page 107; 158pp; English.
 XX
 CC The present invention describes polynucleotides which encode processing
 CC enzymes (e.g. alpha-amylase, pullulanase, alpha-glucosidase, glucose
 CC isomerase, or glucoamylase) that are optimised for expression in plants.
 CC The polynucleotides encode mesophilic, thermophilic or hyperthermophilic
 CC processing enzymes, which are activated under suitable conditions to act
 CC upon the desired substrate. Also described are self-processing transgenic
 CC plants and plant parts, e.g. grain, which express one or more of these
 CC enzymes and have an altered composition that facilitates plant and grain
 CC processing. Also described is a method (M) for converting starch to
 CC starch-derived products in a transformed plant part (TPP), by activating
 CC the starch processing enzyme contained in it. Transgenic grain is useful
 CC for preparing maltodextrin. A transformed plant (TP) can be used to
 CC produce food products having improved taste and to produce fermentable

CC substrates for ethanol and fermented beverages. (M) eliminates the need
 CC to mill or physically disrupt the integrity of plant parts prior to
 CC recovery of starch-derived products. The present sequence represents
 CC alpha-amylase/glucoamylase fusion protein, which is given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 1095 AA;

Query Match 100.0%; Score 78; DB 6; Length 1095;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
 |||||
 Db 78 AYHGYWQDDIYSL 90
 |||||

RESULT 9
 ABB80177
 ID ABB80177 standard; protein; 495 AA.
 XX
 AC ABB80177;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE A. fumigatus AFAA11.
 XX
 KW Enzyme; tannase; cellulase; glucose oxidase; glucoamylase; phytase;
 KW beta-galactosidase; invertase; lipase; alpha-amylase; laccase;
 KW polygalacturonase; xylanase; gallate ester linkage; detergent; cellulose;
 KW glucose; oxygen; myo-inositol phosphate; lactose; tea leaf; sucrose;
 KW glyceride; starch; maltodextrin; oxidated phenolic compound;
 KW polygalacturonic acid chain; xylan; xylo-oligomer; food; feed; beverage;
 KW textile; tea liquor; cleaning ability.
 XX
 OS *Aspergillus fumigatus*.
 XX
 PN WO2003012071-A2.
 XX
 PD 13-FEB-2003.
 XX
 PF 05-AUG-2002; 2002WO-US024842.
 XX
 PR 03-AUG-2001; 2001US-0309870P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Jiang B, Storms R, Roemer T, Bussey H;
 XX
 DR WPI; 2003-332729/31.
 DR N-PSDB; ABQ80345, ABQ80346.
 XX
 PT Novel isolated *Aspergillus fumigatus* polypeptide, useful in various
 PT industries such as those involved in the making of food and feed,
 PT beverages, textiles and detergents.
 XX
 PS Claim 17; Page 134-35; 169pp; English.
 XX
 CC The sequences given in ABB80164-87 show enzymatic proteins derived from
 CC A. fumigatus. These proteins display the catalytic activity of an enzyme
 CC such as tannase, cellulase, glucose oxidase, glucoamylase, phytase, beta-
 CC galactosidase, invertase, lipase, alpha-amylase, laccase,
 CC polygalacturonase or xylanase. Compositions comprising the tannase are
 CC useful for modulating the amount of compounds that comprise a gallate
 CC ester linkage in a composition. Compositions comprising cellulase are
 CC useful for modulating the amount of cellulose in a composition.
 CC Compositions comprising glucose oxidase are useful for modulating the
 CC amount of glucose or oxygen in a composition. Compositions comprising
 CC phytase are useful for modulating the amount of myo-inositol phosphates
 CC in a composition. Compositions comprising beta-galactosidase are useful
 CC for modulating the amount of lactose in a composition. Compositions
 CC comprising sucrose or invertase are useful for modulating the amount of
 CC sucrose in a composition. Compositions comprising lipase are useful for

modulating the amount of glyceride in a composition. Compositions comprising alpha-amylases are useful for modulating the amount of starches or maltodextrins in a composition. Compositions comprising lactase are useful for modulating the amount of oxidated phenolic compounds in a composition. Compositions comprising polygalacturonases are useful for modulating the amount of high or low molecular weight polygalacturonic acid chains in a composition. Compositions comprising xylanases are useful for modulating the amount of xylan or xylo-oligomers in a composition. The *A. fumigatus* proteins and corresponding DNA's are useful in various industries such as those involved in the making of food and feed, beverages, textiles and detergents. The DNA's are useful to express recombinant enzymes for characterization, modification or industrial uses, to compare with the nucleotide sequence of *A. fumigatus* to identify duplicated genes of paralogs having the same or similar biochemical activity and/or function, to compare with nucleic acid sequence of other related or distant fungal organisms to identify potential orthologous enzyme genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, and to raise anti-protein antibodies. The polypeptide having tannase activity increases the yield of tea liquor from tea leaves, improves the colour, flavour and health benefits of tea products, particularly an instant tea product. The polypeptide having cellulase activity enhances cleaning ability of detergent compositions

XX
SQ Sequence 495 AA;

Query Match 96.2%; Score 75; DB 6; Length 495;
Best Local Similarity 92.3%; Pred. No. 0.0032; 0; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 0;

QY 1 AYHGYWQDDIYSL 13
| | | | | | | | | |
Db 97 AYHGYWQDDIYSV 109

RESULT 10
ABB09072
ID ABB09072 standard; protein; 423 AA.
AC ABB09072;
XX
XX
DT 26-JUN-2002 (first entry)
XX
DE Aspergillus oryzae TAKA protein (TAA).
XX
XX Thermus sp. IM6501; maltogenic amylase; EC 3.2.1.133; crystallisation;
KW protein co-ordinate data; 3 dimensional structure.
XX
XX Aspergillus oryzae.

XX KR2001027418-A.
XX
XX PD 06-APR-2001.
XX
XX PF 09-SEP-1999; 99KR-00039130.
XX
XX PR 09-SEP-1999; 99KR-00039130.
XX
XX PA (POST-) POSTECH FOUND.
XX (SAMY-) SAMYANG GENEX CORP.
XX
XX PI Kim TJ, Park GH;
XX
XX DR WPI; 2001-534477/59.
XX
XX Manufacturing maltogenic amylase having improved transglycosylation activity, comprises using crystallization.

XX Disclosure; Page 188; 196pp; Korean.
XX
XX The present invention describes manufacturing maltogenic amylase (EC 3.2.1.133) having improved transglycosylation activity, comprising using crystallisation and the three dimensional structure of maltogenic

CC amylase. Manufacturing maltogenic amylase comprises the following steps:
CC (i) obtaining a gene of maltogenic amylase from *Thermus* sp. IM6501 (KCTC 5027BP) and inserting the gene into plasmid pUC119 to construct recombinant DNA (pThMA119); (ii) inserting the recombinant DNA to *Escherichia coli* MC1061, which is cultivated at 37 plus degrees Celsius for 10 hours in Luria-Bertani (LB) media and centrifuging the media to obtain a microbial cell; (iii) suspending the microbial cell with buffer solution at pH 7.5 and obtaining supernatant; and (iv) passing the supernatant through column chromatography and obtaining purified maltogenic amylase. The maltogenic amylase is a dimer comprised of two maltogenic amylase molecules, and a *Thermus* sp. IM6501 maltogenic amylase (ThMA) crystal. The amylase has a structure containing an activated region that consists of amino acid residues of Asp-328, Glu-357, Asp-424, and a pocket with glucose bound that consists of amino acid residues of Pro44, Tyr45, Arg81, Arg83, Pro118, Cys116, Asn131, Glu132, Val1329, CC and His1360. The present sequence represents *Aspergillus oryzae* TAKA protein (TAA), given in comparison with ThMA in the present invention
XX
SQ Sequence 423 AA;

Query Match 89.7%; Score 70; DB 4; Length 423;
Best Local Similarity 92.3%; Pred. No. 0.016; 1; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 1;

QY 1 AYHGYWQDDIYSL 13
| | | | | | | | | |
Db 74 AMHGYWQDDIYSL 86

RESULT 11
ABP97894
ID ABP97894 standard; protein; 494 AA.
XX
XX AC ABP97894;
XX
XX DT 17-JUN-2003 (first entry)
XX

DE Amino acid sequence of an alpha-amylase of *Aspergillus niger*.

XX Alpha-amylase; enzyme; endohydrolysis; 1,4-alpha-glucosidic linkage;
KW oligosaccharide; polysaccharide; baking.

XX OS Aspergillus niger.

XX PN WO2003016535-A2.

XX PD 27-FEB-2003.

XX PF 02-AUG-2002; 2002WO-NL000522.

XX PR 16-AUG-2001; 2001EP-00000379.

XX PR 16-AUG-2001; 2001EP-00000380.

XX PR 16-AUG-2001; 2001EP-00000381.

XX PR 16-AUG-2001; 2001EP-00000382.

XX PR 16-AUG-2001; 2001EP-00000383.

XX PR 16-AUG-2001; 2001EP-00000384.

XX PA (STAM) DSM NV.

XX MAier D, Stock A, Wagner C, Folkers U, Albermann K, Hopper S;

XX WPI; 2003-312758/30.

XX New polynucleotide from a filamentous fungus, preferably *Aspergillus niger*, useful in a baking process.

XX PS Claim 13; Page 61-63; 81pp; English.

XX The present sequence is an alpha-amylase enzyme of the filamentous fungus *Aspergillus niger*. Alpha-amylases catalyse the endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. The alpha-amylase polynucleotides and polypeptides of the invention are useful in a baking process


```

XX      Sequence 494 AA;
SQ
Query Match      89.7%; Score 70; DB 6; Length 494;
Best Local Similarity 84.6%; Pred. No. 0.019;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
      :|||||:|:|
DB      94 AYHGYWQDAYAL 106

RESULT 12
AAR8213
ID      AAR8213 standard; peptide; 55 AA.
XX
AC      AAR8213;
XX
DT      16-OCT-2003 (revised)
DT      03-APR-1996 (first entry)
XX
DE      Alpha-amylase fragment.
XX
KW      Alpha-amylase; thermostable enzyme; baking; Thermomyces lanuginosus; PCR;
KW      polymerase chain reaction.
XX
OS      Thermomyces lanuginosus; CBS 223.63.
XX
PN      WO9601323-A1.
XX
PD      18-JAN-1996.
XX
PF      03-JUL-1995; 95WO-EP002607.
XX
PR      04-JUL-1994; 94GB-00013419.
XX
PA      (DANI-) DANISCO AS.
XX
PI      Michelsen B, Rasmussen P;
DR      WPI; 1996-087673/09.
DR      N-PSDB; AAT10565.
XX
Thermophilic alpha-amylase with activity range of 60-80 degrees C -
derived from Thermomyces lanuginosus, useful in the prepn. of foodstuffs
and bakery prods. esp. bread.
XX
PS      Claim 3; Page 36-38; 94pp; English.
XX
CC      A thermostable alpha-amylase (AAR8213) of Thermomyces lanuginosus CBS
CC      224.63 is expressed in Aspergillus niger hosts using a gene (AAT10562)
CC      isolated from a T. lanuginosus gene library. The recombinant enzyme (54-
CC      60 kDa) shows optimal activity at 60-70 deg and pH 5.8-6, has a pI of 3.7
CC      and is active at 60-80 deg. (Updated on 16-OCT-2003 to standardise OS
CC      field)
XX
SQ      Sequence 493 AA;

Query Match      88.5%; Score 69; DB 2; Length 493;
Best Local Similarity 76.9%; Pred. No. 0.027;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
      :|||||:|:|
DB      98 SYHGYWQKDLIYL 110

RESULT 14
ADS23436
ID      ADS23436 standard; protein; 500 AA.
XX
AC      ADS23436;
XX
DT      02-DEC-2004 (first entry)
XX
DE      Bacterial polypeptide #12469.
XX
KW      Recombinant DNA construct; transformed plant; improved plant property;
KW      cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW      pathogen tolerance; pest tolerance; plant disease resistance;
KW      cell cycle pathway modification; plant growth regulator;
KW      homologous recombination; seed oil yield; protein yield; carbohydrate;
KW      nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW      bacterial polypeptide.
XX
OS      Bacteria.
XX
PN      US2003233675-A1.
XX

```

PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 12469; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 500 AA;
SQ
Query Match 88.5%; Score 69; DB 8; Length 500;
Best Local Similarity 84.6%; Pred. No. 0.027;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AYHGYYQQDIYSL 13
Db 82 AYHGYYQQDIYSL 94
RESULT 15
ADN19570
ID ADN19570 standard; protein; 513 AA.
XX
XX ADN19570;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #2223.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2223; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 513 AA;
SQ
Query Match 85.9%; Score 67; DB 8; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.057;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 AYHGYYQQDIYSL 13
Db 100 AYHGYYQQDIYSL 112

Search completed: February 8, 2005, 18:36:39
Job time : 7.787 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:29:19 ; Search time 2.00772 Seconds
(without alignments)
483.353 Million cell updates/sec

Title: us-10-820-200-2_COPY_98_110

Perfect score: 78

Sequence: 1 AYHGYQQDIYSL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCFUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfilee1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	478	1	US-08-720-899-7
2	78	100.0	478	1	US-08-459-610-7
3	78	100.0	478	2	US-08-343-804-7
4	78	100.0	478	3	US-08-600-908A-10
5	78	100.0	478	3	US-08-683-838A-10
6	78	100.0	478	3	US-09-182-859-7
7	78	100.0	478	4	US-09-672-459-7
8	78	100.0	478	4	US-09-636-252A-10
9	78	100.0	478	4	US-10-186-042-7
10	62	79.5	478	2	US-08-339-715A-2
11	54	69.2	468	1	US-08-204-656B-2
12	54	69.2	468	1	US-08-470-702-6
13	54	69.2	468	1	US-08-467-831-6
14	53	67.9	468	1	US-08-204-656B-4
15	53	67.9	468	1	US-08-470-702-7
16	53	67.9	468	1	US-08-467-831-7
17	50	64.1	468	1	US-08-204-656B-6
18	50	64.1	468	1	US-08-470-702-8
19	50	64.1	468	1	US-08-467-831-8
20	49	62.8	468	1	US-08-204-656B-8
21	49	62.8	468	1	US-08-470-702-9
22	49	62.8	468	1	US-08-467-831-9
23	49	62.8	685	3	US-08-947-965-72
24	49	62.8	685	3	US-08-947-965-74
25	49	62.8	686	3	US-08-947-965-70
26	49	62.8	686	3	US-08-947-965-73
27	47	60.3	634	3	US-08-947-965-78

28 47 60.3 655 1 US-08-469-202-27 Sequence 27, Appl
29 47 60.3 655 1 US-08-469-202-28 Sequence 28, Appl
30 47 60.3 655 2 US-08-484-434C-34 Sequence 34, Appl
31 47 60.3 655 2 US-08-484-434C-35 Sequence 35, Appl
32 47 60.3 655 4 US-09-384-361-34 Sequence 34, Appl
33 47 60.3 655 4 US-09-384-361-35 Sequence 35, Appl
34 46 59.0 675 3 US-08-947-965-76 Sequence 76, Appl
35 46 59.0 676 3 US-08-947-965-71 Sequence 71, Appl
36 46 59.0 680 3 US-08-947-965-77 Sequence 77, Appl
37 46 59.0 683 3 US-08-947-965-2 Sequence 2, Appl
38 46 59.0 687 3 US-08-947-965-75 Sequence 75, Appl
39 46 59.0 725 2 US-08-816-105A-1 Sequence 1, Appl
40 45 57.7 719 3 US-09-386-607-2 Sequence 2, Appl
41 45 57.7 719 4 US-09-645-707B-2 Sequence 2, Appl
42 44 56.4 14 1 US-08-204-656B-11 Sequence 11, Appl
43 44 56.4 14 1 US-08-470-702-11 Sequence 11, Appl
44 44 56.4 14 1 US-08-467-831-11 Sequence 11, Appl
45 44 56.4 454 1 US-07-930-686-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 100.0%; Score 78; DB 1; Length 478;
Best Local Similarity 100.0%; Pred No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 5801043o No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match      100.0%; Score 78; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 4
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 5989169o No. 5989169disk of No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; US-08-600-908A-10

Query Match      100.0%; Score 78; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 5
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 5989169o No. 5989169disk of No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; US-08-600-908A-10
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-600-908A-10

Query Match      100.0%; Score 78; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 5
US-08-683-838A-10
; Sequence 10, Application US/08693838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: "-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6022724 No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,838A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-683-838A-10

Query Match      100.0%; Score 78; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 6
US-09-182-859-7
; Sequence 7, Application US/09182859
; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisg rd-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match      100.0%; Score 78; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 7
US-09-672-459-7
; Sequence 7, Application US/09672459
; Patent No. 6436888
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisg rd-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/672,459
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-672-459-7

Query Match      100.0%; Score 78; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90
```

RESULT 8

US-09-636-252A-10
; Sequence 10, Application US/09636252A
; Patent No. 6440716
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/1f216-US2
; CURRENT APPLICATION NUMBER: US/09/636,252A
; CURRENT FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683,838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 478
; TYPE: PRT
; ORGANISM: A. oryzae
US-09-636-252A-10

Query Match 100.0%; Score 78; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
||| ||||| |||||
Db 78 AYHGYWQDIYSL 90

RESULT 9

US-10-186-042-7
; Sequence 7, Application US/10186042
; Patent No. 6642044
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/186,042
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/672,459
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-186-042-7

Query Match 100.0%; Score 78; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
||| ||||| |||||
Db 78 AYHGYWQDIYSL 90

RESULT 10

US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306096/1993
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Melser, Allen S.
; REGISTRATION NUMBER: 27,215
; REFERENCE/DOCKET NUMBER: 18335.009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 79.5%; Score 62; DB 2; Length 478;
Best Local Similarity 84.6%; Pred. No. 0.044;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
||| ||||| |||||
Db 78 AYGYWQDIYSL 90

RESULT 11

US-08-204-656B-2
; Sequence 2, Application US/08204656B
; Patent No. 5538892
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme

```

; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S..
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-470-702-6
; Query Match 69.2%; Score 54; DB 1; Length 468;
; Best Local Similarity 53.8%; Pred. No. 0.82;
; Matches 7; Conservative 4; Mismatches 2; Indels
;
; QY 1 AYHGYWQDDIYL 13
; |||||::|||:
; Db 79 AYHGFWMKNIKY 91
;
; RESULT 13
; US-08-467-831-6
; Sequence 6, Application US/08467831

```

0: Gaps 0:

APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD

CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-467-831-6

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.82;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
Db 79 AYHGWMMKNIYKI 91

RESULT 14

US-08-204-656B-4
Sequence 4, Application US/08204656B

Patent No. 5538882

GENERAL INFORMATION:

APPLICANT: Matsui, Ikuo
APPLICANT: Ishikawa, Kazuhiko

APPLICANT: Miyairi, Sachio

APPLICANT: Honda, Koichi

TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,

TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing

TITLE OF INVENTION: Oligosaccharide Using The Enzyme

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP

STREET: 8110 Gatehouse Road, Suite 500 East

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/204,656B

FILING DATE: 02-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Weiner, Marc S.

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 234-252P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-204-656B-4

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 1.2;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
Db 79 AYHGWMMKNIYKI 91

RESULT 15

US-08-470-702-7

Sequence 7, Application US/08470702

Patent No. 5631149

GENERAL INFORMATION:

APPLICANT: MATSUI, IKUO

APPLICANT: ISHIKAWA, KAZUHIKO

APPLICANT: MIYAIRI, SACHIO

APPLICANT: HONDA, KOICHI

TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,

TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

STREET: 8110 Gatehouse Road, Suite 500 East

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,702

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/204,656

FILING DATE: 02-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: WEINER, MARC S.

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 234-252P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-470-702-7

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 1.2;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
Db 79 AYHGWMMKNIYKI 91

Search completed: February 8, 2005, 18:44:37
Job time : 2.67439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:42:07 ; Search time 5.59653 Seconds
(without alignments)
756.564 Million cell updates/sec

Title: US-10-820-200-2_copy_98_110

Perfect score: 78

Sequence: 1 AYHGYYQDIYSL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	478	14	US-10-184-771-10 Sequence 10, Appl
2	78	100.0	478	14	US-10-186-042-7 Sequence 7, Appl
3	78	100.0	478	15	US-10-644-187-7 Sequence 7, Appl
4	78	100.0	478	17	US-10-926-720-10 Sequence 10, Appl
5	78	100.0	498	16	US-10-815-495-22 Sequence 22, Appl
6	78	100.0	499	16	US-10-815-495-18 Sequence 18, Appl
7	78	100.0	1095	14	US-10-228-063-45 Sequence 45, Appl
8	75	96.2	495	14	US-10-213-990-42 Sequence 42, Appl
9	69	88.5	500	15	US-10-369-493-12469 Sequence 12469, A
10	67	85.9	513	15	US-10-369-493-2223 Sequence 2223, Ap
11	65	83.3	547	14	US-10-213-990-48 Sequence 48, Appl
12	64	82.1	484	15	US-10-416-393-1 Sequence 1, Appl
13	58	74.4	630	14	US-10-213-990-45 Sequence 45, Appl

14	56	71.8	493	15	US-10-369-493-22723	Sequence 22723, A
15	53	67.9	564	15	US-10-369-493-22679	Sequence 22679, A
16	49	62.8	686	16	US-10-872-198-44	Sequence 44, Appl
17	48	61.5	502	14	US-10-081-872-108	Sequence 108, App
18	48	61.5	502	15	US-10-385-305-108	Sequence 108, App
19	47	60.3	483	15	US-10-369-493-597	Sequence 597, App
20	46	59.0	170	10	US-09-791-932-107	Sequence 107, App
21	46	59.0	530	15	US-10-369-493-15914	Sequence 15914, A
22	46	59.0	570	10	US-09-927-827-45	Sequence 45, Appl
23	46	59.0	702	14	US-10-188-066-2	Sequence 2, Appl
24	46	59.0	713	14	US-10-332-937-2	Sequence 2, Appl
25	45	57.7	557	14	US-10-081-872-206	Sequence 206, App
26	45	57.7	557	15	US-10-385-305-206	Sequence 206, App
27	45	57.7	719	14	US-10-234-266-2	Sequence 2, Appl
28	45	57.7	719	15	US-10-442-558-2	Sequence 2, Appl
29	45	57.7	719	15	US-10-453-828-2	Sequence 2, Appl
30	45	57.7	929	14	US-10-081-872-196	Sequence 196, App
31	45	57.7	929	15	US-10-385-305-196	Sequence 196, App
32	44	56.4	676	15	US-10-369-493-23599	Sequence 23599, A
33	44	56.4	833	13	US-10-014-436-3	Sequence 3, Appl
34	44	56.4	1938	13	US-10-014-436-2	Sequence 2, Appl
35	44	56.4	2022	16	US-10-408-765A-598	Sequence 598, App
36	43	55.1	640	10	US-09-769-736-123	Sequence 123, App
37	43	55.1	747	9	US-09-978-295A-459	Sequence 459, App
38	43	55.1	747	9	US-09-938-418-9	Sequence 9, Appl
39	43	55.1	747	9	US-09-978-697-459	Sequence 459, App
40	43	55.1	747	9	US-09-978-192A-459	Sequence 459, App
41	43	55.1	747	9	US-09-999-832A-459	Sequence 459, App
42	43	55.1	747	10	US-09-978-189-459	Sequence 459, App
43	43	55.1	747	10	US-09-978-608A-459	Sequence 459, App
44	43	55.1	747	10	US-09-978-585A-459	Sequence 459, App
45	43	55.1	747	10	US-09-978-191A-459	Sequence 459, App

ALIGNMENTS

RESULT 1
US-10-184-771-10
; Sequence 10, Application US/10184771
; Publication No. US20030170769A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/15216-US2
; CURRENT APPLICATION NUMBER: US/10/184,771
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/636,252
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683,838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 478
; TYPE: PRT
; ORGANISM: A. oryzae
US-10-184-771-10

Query Match 100.0%; Score 78; DB 14; Length 478;
Best Local Similarity 100.0%; Pred. NO. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AYHGYYQDIYSL 13

Db 78 AYHGYYQDIYSL 90

RESULT 2

US-10-186-042-7
; Sequence 7, Application US/10186042

Publication No. US20030171236A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/186,042
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/672,459
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-186-042-7

Query Match 100.0%; Score 78; DB 14; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
DB 78 AYHGYWQDIYSL 90

RESULT 3
US-10-644-187-7
; Sequence 7, Application US/10644187
; Publication No. US20040048351A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/644,187
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-644-187-7

Query Match 100.0%; Score 78; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13

DB 78 AYHGYWQDIYSL 90

RESULT 4

US-10-926-720-10
; Sequence 10, Application US/10926720
; Publication No. US20050019886A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: -Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/926,720
; FILING DATE: 26-Aug-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-926-720-10

Query Match 100.0%; Score 78; DB 17; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
DB 78 AYHGYWQDIYSL 90

RESULT 5

US-10-815-495-22
; Sequence 22, Application US/10815495
; Publication No. US20040191864A1
; GENERAL INFORMATION:
; APPLICANT: Novozymes Biotech, Inc.
; APPLICANT: Connelly, Mariah
; APPLICANT: Brody, Howard
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient
; TITLE OF INVENTION: Mutants Of Aspergillus Niger
; FILE REFERENCE: 10345.200-US
; CURRENT APPLICATION NUMBER: US/10/815,495
; CURRENT FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 30

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-815-495-22

Query Match      100.0%; Score 78; DB 16; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      99 AYHGYWQDDIYSL 111

RESULT 6
US-10-815-495-18
; Sequence 18, Application US/10815495
; Publication No. US20040191864A1
; GENERAL INFORMATION:
; APPLICANT: Novozymes Biotech, Inc.
; APPLICANT: Brody, Howard
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient
; TITLE OF INVENTION: Mutants Of Aspergillus Niger
; FILE REFERENCE: 10345.200-US
; CURRENT APPLICATION NUMBER: US/10/815,495
; CURRENT FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-815-495-18

Query Match      100.0%; Score 78; DB 16; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      99 AYHGYWQDDIYSL 111

RESULT 7
US-10-228-063-45
; Sequence 45, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Aspergillus shirousami
US-10-228-063-45

Query Match      100.0%; Score 78; DB 14; Length 1095;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      78 AYHGYWQDDIYSL 90

RESULT 8
US-10-213-990-42
; Sequence 42, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Aspergillus
US-10-213-990-42

Query Match      96.2%; Score 75; DB 14; Length 495;
Best Local Similarity 92.3%; Pred. No. 0.0026;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      97 AYHGYWQDDIYSL 109

RESULT 9
US-10-369-493-12469
; Sequence 12469, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12469
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-10-369-493-12469

Query Match      88.5%; Score 69; DB 15; Length 500;
Best Local Similarity 84.6%; Pred. No. 0.021;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AYHGYWQDDIYSL 13
Db      82 AYHGYWQDDIYSL 94

RESULT 10
US-10-369-493-2223
; Sequence 2223, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
```

APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2223
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2223

Query Match 85.9%; Score 67; DB 15; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.044; 1; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :|:
Db 100 AYHGYWQDDIYTL 112

RESULT 11
US-10-213-990-48
; Sequence 48, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Aspergillus
US-10-213-990-48

Query Match 83.3%; Score 65; DB 14; Length 547;
Best Local Similarity 76.9%; Pred. No. 0.095; 1; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :|:
Db 103 AYHGYWQDDIYAL 115

RESULT 12
US-10-416-393-1
; Sequence 1, Application US/10416393
; Publication No. US20040091983A1
; GENERAL INFORMATION:
; APPLICANT: Veit, Christopher
; APPLICANT: Felby, Claus
; APPLICANT: Fugleang, Claus
; TITLE OF INVENTION: SECONDARY LIQUEFACTION IN ETHANOL PRODUCTION
; FILE REFERENCE: 10113.204-US
; CURRENT APPLICATION NUMBER: US/10/416,393
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Aspergillus niger

FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SEQ ID NO:1
US-10-416-393-1

Query Match 82.1%; Score 64; DB 15; Length 484;
Best Local Similarity 76.9%; Pred. No. 0.12; 2; Indels 0; Gaps 0;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :|:
Db 78 AYHGYWQDDIYDV 90

RESULT 13
US-10-213-990-45
; Sequence 45, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Aspergillus
US-10-213-990-45

Query Match 74.4%; Score 58; DB 14; Length 630;
Best Local Similarity 61.5%; Pred. No. 1.3; 4; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :|:
Db 101 SYHGYWQDDIHEV 113

RESULT 14
US-10-369-493-22723
; Sequence 22723, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22723
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(493)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-22723

Query Match 71.8%; Score 56; DB 15; Length 493;
Best Local Similarity 69.2%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
| | | | | | | | | |
Db 89 AYHGYWQDDMTQL 101

RESULT 15
US-10-369-493-22679
; Sequence 22679, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22679
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-22679

Query Match 67.9%; Score 53; DB 15; Length 564;
Best Local Similarity 69.2%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
| | | | | | | | | |
Db 100 AYHGYWNTDYESL 112

Search completed: February 8, 2005, 19:05:53
Job time : 6.92986 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:28:34 ; Search time 1.50579 Seconds
(without alignments)
830.671 Million cell updates/sec

Title: US-10-820-200-2_COPY_98_110

Perfect score: 78

Sequence: 1 AVHGYWQDIYSL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	498	2 A48305	alpha-amylase (EC 3.2.1.1)
2	78	100.0	499	1 ALAS1	alpha-amylase (EC 3.2.1.1)
3	78	100.0	499	1 ALAS3	alpha-amylase (EC 3.2.1.1)
4	78	100.0	499	2 JS0663	alpha-amylase (EC 3.2.1.1)
5	78	100.0	499	2 JT0466	alpha-amylase (EC 3.2.1.1)
6	78	100.0	499	2 JN0588	alpha-amylase (EC 3.2.1.1)
7	78	100.0	499	2 B48305	alpha-amylase (EC 3.2.1.1)
8	67	85.9	513	2 T38770	alpha-amylase (EC 3.2.1.1)
9	66	84.6	631	2 S72270	alpha-amylase (EC 3.2.1.1)
10	64	82.1	484	1 A35282	alpha-amylase (EC 3.2.1.1)
11	64	82.1	507	2 J33921	alpha-amylase (EC 3.2.1.1)
12	62	79.5	478	2 JK0201	alpha-amylase (EC 3.2.1.1)
13	62	79.5	624	1 JC4510	pullulanase (EC 3.2.1.1)
14	59	75.6	482	2 S31478	alpha-amylase (EC 3.2.1.1)
15	58	74.4	494	1 ALBYAF	alpha-amylase (EC 3.2.1.1)
16	56	71.8	478	2 T40860	probable alpha-amylase (EC 3.2.1.1)
17	54	69.2	491	2 T38448	probable alpha-amylase (EC 3.2.1.1)
18	53	67.9	554	2 T41503	probable alpha-amylase (EC 3.2.1.1)
19	52	66.7	581	2 S62505	probable alpha-amylase (EC 3.2.1.1)
20	50	64.1	512	2 S23355	alpha-amylase (EC 3.2.1.1)
21	50	64.1	512	2 S06115	alpha-amylase (EC 3.2.1.1)
22	50	64.1	625	2 T41603	alpha-amylase (EC 3.2.1.1)
23	49	62.8	528	1 ALBSK	alpha-amylase (EC 3.2.1.1)
24	49	62.8	690	2 B82409	alpha-amylase (EC 3.2.1.1)
25	49	62.8	712	1 ALBSG3	cyclomaltodextrin
26	49	62.8	713	1 ALBSG7	cyclomaltodextrin
27	49	62.8	713	1 ALBSG1	cyclomaltodextrin
28	49	62.8	713	1 ALBSXR	cyclomaltodextrin
29	49	62.8	713	2 A58800	cyclomaltodextrin

30	49	62.8	713	2 S09196	cyclomaltodextrin
31	49	62.8	718	1 ALBSG6	cyclomaltodextrin
32	49	62.8	718	1 ALBSGC	cyclomaltodextrin
33	49	62.8	718	1 ALBSMX	cyclomaltodextrin
34	47	60.3	483	2 G75392	glycosyl hydrolase
35	47	60.3	655	1 ALKBG	cyclomaltodextrin
36	46	59.0	703	1 ALBSX1	cyclomaltodextrin
37	46	59.0	704	2 I39805	cyclomaltodextrin
38	46	59.0	710	2 S63598	cyclomaltodextrin
39	46	59.0	711	1 ALBSXF	cyclomaltodextrin
40	46	59.0	714	1 ALBSGR	cyclomaltodextrin
41	45	57.7	717	1 S28784	alpha-amylase (EC 3.2.1.1)
42	44	56.4	675	2 AH0979	alpha-amylase (EC 3.2.1.1)
43	44	56.4	676	2 F91185	alpha-amylase (EC 3.2.1.1)
44	44	56.4	676	2 D86032	alpha-amylase (EC 3.2.1.1)
45	44	56.4	676	2 S23807	alpha-amylase (EC 3.2.1.1)

ALIGNMENTS

RESULT 1

A48305

alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C:Accession: A48305

R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;

Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus

A:Reference number: A48305; MUID:90254827; PMID:2340591

A:Accession: A48305

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-498 <KOR>

A:Cross-references: UNIPROT:Q02905

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 78; DB 2; Length 498;

Best Local Similarity 100.0%; Pred. No. 4.1e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVHGYWQDIYSL 13

DB 99 AVHGYWQDIYSL 111

RESULT 2

ALAS1

alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C:Accession: S04548; A33214; JS0240; A91930; A93767; A10627

R:Wirsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.

Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon organization

A:Reference number: S04548; MUID:89237897; PMID:2785629

A:Accession: S04548

A:Molecule type: DNA

A:Residues: 1-499 <WIR>

A:Cross-references: UNIPROT:P10529; EMBL:X12725; NID:G2430; PIDN:CAA31218.1; PID:G295921

A:Genetics: AMY1

A:Molecule type: mRNA

A:Accession: A33214

A:Residues: 1-499 <WIR>

A:Cross-references: GB:X12725; NID:G2430; PIDN:CAA31218.1; PID:G295921

R:Genes, M.J.; Dove, M.J.; Selig, V.L.

Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A:Reference number: JS0240; MUID:89378767; PMID:2789162
A:Accession: JS0240
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Genetics: AMY2
A:Note: the authors refer to this as isozyme II
R:Isemura, S.; Ikenaka, T.
J. Biochem. 74, 1-10, 1973
A:Reference number: A91930; MUID:74001521; PMID:4733850
A:Accession: A91930
A:Molecule type: protein
A:Residues: 206-225 <ISE>
R:Narita, K.
Proc. Jpn. Acad. 51, 285-290, 1975
A:Reference number: A93767
A:Accession: A93767
A:Molecule type: protein
A:Residues: 434-443, 446-447, 'Q', 449-458, 'GTV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>
R:Matsumura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370; PMID:6609921
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkmenburg, J.P.; Wilkinson, A.
submitted to the Brookhaven Protein Data Bank, August 1992
A:Reference number: A51548; PDB:6TAA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics: <AMY1>
A:Gene: amy1
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Genetics: <AMY2>
A:Gene: amy2; AmyII
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polysaccharide
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase 1 #status experimental <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
|||||
DB 99 AYHGYWQDDIYSL 111

RESULT 3
ALIAS
alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S04549; A33215; A44713
R:Wiseel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon c
A:Reference number: S04548; MUID:89237897; PMID:2785629
A:Accession: S04549
A:Molecule type: DNA
A:Residues: 1-499 <WIR>
A:Cross-references: UNIPROT:Q96TH4; EMBL:X12727; MUID:g2454; PIDN:CAA31220.1; PID:g295922

A:Accession: A33215
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-499 <W12>
A:Cross-references: GB:X12727; MUID:g2454; PIDN:CAA31220.1; PID:g295922
R:Genes, M.J.; Dove, M.J.; Selig, V.L.
Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A:Reference number: JS0240; MUID:89378767; PMID:2789162
A:Accession: A44713
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Note: the authors refer to this as isozyme I
R:Matsumura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370; PMID:6609921
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Gene: amy3; AmyI
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; metal
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase 3 #status experimental <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
|||||
DB 99 AYHGYWQDDIYSL 111

RESULT 4
JS0663
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: JS0663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biosci. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousanii and its expression
A:Reference number: JS0663; MUID:92323146; PMID:1368777
A:Accession: JS0663
A:Molecule type: mRNA
A:Residues: 1-499 <SHI>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <ALP>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
|||||

Db 99 AYHGYWQDIYSL 111

RESULT 5
JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus oryzae
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
A:Cross-references: UNIPROT:P10529
A:Comment: See also PIR:JK0201 and PIR:JS0240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 99 AYHGYWQDIYSL 111

RESULT 6
JN0588
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Uda, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for multiple genes
A:Reference number: JN0588; MUID:90128276; PMID:2612911
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1-499 <TSU>
A:Cross-references: UNIPROT:Q96TH4
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 99 AYHGYWQDIYSL 111

RESULT 7
B48305
alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.; et al.
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus awamori
A:Reference number: A48305; MUID:90254827; PMID:2340591
A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
A:Cross-references: UNIPROT:Q02906
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 99 AYHGYWQDIYSL 111

RESULT 8
T38770
alpha-amylase a precursor - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38770
R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21751
A:Accession: T38770
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-513 <SKE>
A:Cross-references: UNIPROT:O14154; EMBL:Z98762; PIDN:CAB11471.1; GSPDB:GN00066; SPDB:SP7
A:Experimental source: strain 972h-; cosmid c4A8
C:Genetics:
A:Gene: SPDB:SPAC4A8.01
A:Map position: 1
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

Query Match 85.9%; Score 67; DB 2; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.0027;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 100 AYHGYWQDIYSL 112

RESULT 9
S72270
alpha-amylase (EC 3.2.1.1) precursor - Cryptococcus sp. (strain CS2)
C:Species: Cryptococcus sp.
A:Variety: strain CS2
C>Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 99 AYHGYWQDIYSL 111

RESULT 7
B48305
alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.; et al.
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus awamori
A:Reference number: A48305; MUID:90254827; PMID:2340591
A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
A:Cross-references: UNIPROT:Q02906
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 99 AYHGYWQDIYSL 111

RESULT 8
T38770
alpha-amylase a precursor - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38770
R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21751
A:Accession: T38770
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-513 <SKE>
A:Cross-references: UNIPROT:O14154; EMBL:Z98762; PIDN:CAB11471.1; GSPDB:GN00066; SPDB:SP7
A:Experimental source: strain 972h-; cosmid c4A8
C:Genetics:
A:Gene: SPDB:SPAC4A8.01
A:Map position: 1
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

Query Match 85.9%; Score 67; DB 2; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.0027;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
Db 100 AYHGYWQDIYSL 112

RESULT 9
S72270
alpha-amylase (EC 3.2.1.1) precursor - Cryptococcus sp. (strain CS2)
C:Species: Cryptococcus sp.
A:Variety: strain CS2
C>Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C;Accession: S72270
R;Refuji, H.; Chino, M.; Kato, M.; Imura, Y.
Biochem. J. 318, 989-996, 1996
A;Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus*
A;Reference number: S72270; MUID:96433120; PMID:8836148
A;Accession: S72270
A;Molecule type: DNA
A;Residues: 1-631 <EF>
A;Cross-references: UNIPROT:Q92394; EMBL:D83540; NID:g1595852; PIDN:BAA12010.1; PID:g1595852
A;Experimental source: strain S-2
C;Genetics:
A;Gene: amy-CS2
A;Introns: 289/2; 326/2
C;Function:
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-631/Product: alpha-amylase #status predicted <MAT>
F;206-335/Domain: alpha-amylase core homology <AMY>

Query Match 84.6%; Score 66; DB 2; Length 631;
Best Local Similarity 76.9%; Pred. No. 0.005;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGYWQQDIYSL 13
| | | | |
Db 105 AYHGYWAQDIYEI 117

RESULT 10
A35282
alpha-amylase (EC 3.2.1.1) - *Aspergillus niger*
C;Species: *Aspergillus niger*
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35282
R;Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; Petersen
Biochemistry 29, 6244-6249, 1990
A;Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom resolution
A;Reference number: A35282; MUID:91002514; PMID:2207069
A;Accession: A35282
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-484 <BOE>
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;173-300/Domain: alpha-amylase core homology <AMY>

Query Match 82.1%; Score 64; DB 1; Length 484;
Best Local Similarity 76.9%; Pred. No. 0.008;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGYWQQDIYSL 13
| | | | |
Db 78 AYHGYWQKIYDV 90

RESULT 11
S33921
alpha-amylase (EC 3.2.1.1) SWA2 precursor - yeast (*Schwanniomyces occidentalis*)
N;Alternate names: alpha-1,4 glucanohydrolase
C;Species: *Schwanniomyces occidentalis*
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S33921
R;Claros, M.G.; Abarca, D.; Fernandez-Lobato, M.; Jimenez, A.
Curr. Genet. 24, 75-83, 1993
A;Title: Molecular structure of the SWA2 gene encoding an AMY1-related alpha-amylase from
A;Reference number: S33921; MUID:93365041; PMID:8358835
A;Accession: S33921
A;Molecule type: DNA

A:Residues: 1-507 <CIA>
A:Cross-references: UNIPROT:Q08806; EMBL:X73497; NID:g396561; PIDN:CAA51912.1; PID:g396561
C:Genetics:
A:Gene: SWA2
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-507/Product: alpha-amylase #status predicted <MAT>
F:205-332/Domain: alpha-amylase core homology <AMY>
F:134,229/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 64; DB 2; Length 507;
Best Local Similarity 69.2%; Pred. No. 0.0084;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AYHGYYQDDIYSL 13
| | | | | : | | : |
Db 110 AYHGYYQDDIYAI 122

RESULT 12
JK0201
alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C:Accession: JK0201
R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A:Title: The complete amino acid sequence of Taka-amylase A.
A:Reference number: JK0201
A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOD>
A:Cross-references: UNIPROT:P10529
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:JS0240.
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amylase core homology <AMY>
F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 79.5%; Score 62; DB 2; Length 478;
Best Local Similarity 84.6%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGYYQDDIYSL 13
| | | | | : | | : |
Db 78 AYTGYYQDDIYSL 90

RESULT 13
JC4510
pullulanase (EC 3.2.1.41) precursor - yeast (lipomyces kononenkoae)
N:Alternate names: LXA1 protein; raw starch-degrading amylase
N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
C:Species: Lipomyces kononenkoae
C:Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C:Accession: JC4510; PC4116
R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.
Gene 166, 65-71, 1995
A:Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a Lipon
A:Reference number: JC4510; MUID:96105202; PMID:8528995
A:Accession: JC4510
A:Molecule type: mRNA

A:Residues: 1-624 <STE>
A:Cross-references: UNIPROT:Q01117; GB:U03076; NID:g1173536; PIDN:AAC49622.1; PID:g1173536
A:Experimental source: Strain IGC4052B
A:Accession: PC4116
A:Molecule type: protein
A:Residues: 29-44 <ST2>
A:Experimental source: IGC4052B
C:Genetics:
A:Gene: LKA1
C:Function:
A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
A:Pathway: glycogen/starch degradation
C:Superfamily: Lipomyces alpha-amylase; alpha-amylase core homology; glucoamylase starch
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-624/Product: alpha-amylase #status predicted <MAT>
F:48-141/Domain: glucoamylase starch-binding domain homology <SBD>
F:320-447/Domain: alpha-amylase core homology <AMY>
F:177-185,297-311,387-430,587-622/Dissulfide bonds: #status predicted
F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 79.5%; Score 62; DB 1; Length 624;
Best Local Similarity 69.2%; Pred. No. 0.022;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :||| :|
Db 225 AYHGYWQDDIYSL 237

RESULT 14
S31478
alpha-amylase (EC 3.2.1.1) - Thermoactinomyces vulgaris
C:Species: Thermoactinomyces vulgaris
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S31478
R:Hofemeister, B.; Koenig, S.; Hoang, V.; Engel, J.; Mayer, G.; Heese, O.; Hansen, G.; H
submitted to the EMBL Data Library, December 1992
A:Description: The gene amyTV coding for a non-glucogenic alpha-amylase from Thermoactin
A:Reference number: S31478
A:Accession: S31478
A:Molecule type: DNA
A:Residue: 1-482 <HOF>
A:Cross-references: UNIPROT:Q60051; EMBL:X69807; NID:g48289; PIDN:CAA49465.1; PID:g48290
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:194-319/Domain: alpha-amylase core homology <AMY>

Query Match 75.6%; Score 59; DB 2; Length 482;
Best Local Similarity 69.2%; Pred. No. 0.053;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :||| :|
Db 107 AYHGYWQDDIYSL 119

RESULT 15
ALBYAF
alpha-amylase (EC 3.2.1.1) precursor - yeast (Saccharomycopsis fibuligera)
C:Species: Saccharomycopsis fibuligera
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C:Accession: S00064
R:Itoh, T.; Yamashita, I.; Fukui, S.
FEBS Lett. 219, 339-342, 1987
A:Title: Nucleotide sequence of the alpha-amylase gene (ALP1) in the yeast Saccharomycop
A:Reference number: S00064; MUID:87276512; PMID:3497057
A:Accession: S00064
A:Molecule type: DNA

A:Residues: 1-494 <ITO>
A:Cross-references: UNIPROT:P21567; EMBL:X05791; NID:g4847; PIDN:CAA29233.1; PID:g4848
A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in having 368-Thr
C:Genetics:
A:Gene: ALP1
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; extracellular protein; glycoprotein; glycosidase; hydrolase; metallo
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-494/Product: alpha-amylase #status predicted <MAT>
F:200-327/Domain: alpha-amylase core homology <AMY>
F:57-65,177-191,267-310,462-493/Dissulfide bonds: #status predicted
F:148,189,202,237/Binding site: calcium (Asn, Gln, Asp, His) #status predicted
F:224/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:233,257,324/Active site: Asp, Glu, Asp #status predicted

Query Match 74.4%; Score 58; DB 1; Length 494;
Best Local Similarity 61.5%; Pred. No. 0.079;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
||||| :||| :|
Db 105 AYHGYWQDDIYSL 117

Search completed: February 8, 2005, 18:43:01
Job time : 1.50579 secs

THIS PAGE BLANK (use 12)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:25:28 ; Search time 7.62934 Seconds
(without alignments)
872.556 Million cell updates/sec

Title: US-10-820-200-2_COPY_98_110

Perfect score: 78

Sequence: 1 AYHGYWQDIYSL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	498	1 AMYA ASPAW	Q02905 aspergillus
2	78	100.0	498	2 Q76CT3	Q76ct3 aspergillus
3	78	100.0	499	1 AMYA ASPOR	P10529 aspergillus
4	78	100.0	499	1 AMYB ASPAW	Q02906 aspergillus
5	78	100.0	499	1 AMY ASPSH	P30292 aspergillus
6	78	100.0	499	2 Q96TH4	Q96th4 aspergillus
7	78	100.0	499	2 Q76LJ9	Q76l99 aspergillus
8	78	100.0	499	2 Q9UV07	Q71v45 aspergillus
9	75	96.2	490	2 Q9UV07	Q9uv07 emericella
10	68	87.2	128	2 O15751	O15751 dictyosteli
11	67	85.9	513	1 AMY3 SCHPO	O14154 schizosacch
12	66	84.6	631	2 Q923J4	Q92394 cryptococcu
13	64	82.1	484	1 AMYA ASPNG	P56271 aspergillus
14	64	82.1	507	1 AMY2 DBEOC	Q08806 debaryomyce
15	64	82.1	634	2 Q76LJ9	Q76l99 aspergillus
16	63	80.8	640	2 O13296	O13296 aspergillus
17	62	79.5	624	1 AMY1 LIPKO	Q01117 lipomyces k
18	62	79.5	647	2 Q6VFP3	Q6vfp3 lipomyces s
19	59	75.6	482	2 Q60051	Q60051 thermostactin
20	59	75.6	492	2 Q7SDJ6	Q08806 debaryomyce
21	58	74.4	494	1 AMY1 SACFI	P21567 saccharomyc
22	56	71.8	478	1 YQ29 SCHPO	O10427 schizosacch
23	54	69.2	491	2 O13396	O13396 schizosacch
24	54	69.2	623	2 Q9UV09	Q9uv09 emericella
25	53	67.9	564	1 AMY4 SCHPO	Q9y7s9 schizosacch
26	52	66.7	533	2 Q7S4K0	Q7s4k0 neurospora
27	52	66.7	581	1 AMY1 SCHPO	Q09840 schizosacch
28	52	66.7	711	2 Q992B3	Q992b3 streptococ
29	50	64.1	499	2 Q8J1E4	Q8jie4 lipomyces k
30	50	64.1	512	1 AMY1 DBEOC	P19269 debaryomyce
31	50	64.1	625	2 O74922	O74922 schizosacch

32	49	62.8	528	1 AMY_BACCI	P08137 bacillus ci
33	49	62.8	686	2 O87ER2	O87er2 vibrio para
34	49	62.8	687	2 Q7MCL0	Q7mcl0 vibrio vuln
35	49	62.8	687	2 Q8D5L1	Q8d5l1 vibrio vuln
36	49	62.8	690	2 Q9KL86	Q9kl86 vibrio chol
37	49	62.8	712	1 CDGT_BACS3	P09121 bacillus sp
38	49	62.8	712	2 Q6S3E3	Q6s3e3 bacillus sp
39	49	62.8	713	1 CDG2 PAEMA	P31835 paenibacill
40	49	62.8	713	1 CDGT_BAC11	P30921 bacillus sp
41	49	62.8	713	1 CDGT_BACS0	P05618 bacillus sp
42	49	62.8	713	1 CDGT_BACS8	P17692 bacillus sp
43	49	62.8	713	1 CDGU_BACCI	P43379 bacillus ci
44	49	62.8	713	2 Q9F5W3	Q9f5w3 bacillus ci
45	49	62.8	718	1 CDGT_BACCI	P30920 bacillus ci

ALIGNMENTS

RESULT 1
ID AMYA ASPAW STANDARD; PRT; 498 AA.
AC Q02905;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase A).
GN Name=AMYA;
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L., Kodama K.H.,
Royer T.J., Thompson S.A., Ward M., Wilson L.J., Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori";
RL Curr. Genet. 17:203-212(1990).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52755; CAA36966.1; -
CC PIR; A48305; A48305.
CC HSSP; P10529; 7TAA.
CC InterPro; IPR006589; Alp_amy1_cat_sub.
CC InterPro; IPR006047; Alpha_amy1_cat.
CC InterPro; IPR006046; Glyco_hydro_13.
CC Pfam; PF00128; Alpha-amylase; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Amy; 1.
KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolase; Multigene family; Signal.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 498 Alpha-amylase A.
FT ACT_SITE 227 227 Nucleophile (By similarity).
FT ACT_SITE 251 251 Proton donor (By similarity).
FT ACT_SITE 318 318 By similarity.
FT METAL 142 142 Calcium 1 (By similarity).

```

FT METAL          183      183      Calcium 1 (via carbonyl oxygen) (By
FT METAL          196      196      similarity).
FT METAL          227      227      Calcium 1 (By similarity).
FT METAL          231      231      Calcium 2 (By similarity).
FT METAL          251      251      Calcium 1 (via carbonyl oxygen) (By
FT METAL          251      251      similarity).
FT METAL          51       59       Calcium 2 (By similarity).
FT DISULFID       171      185       By similarity.
FT DISULFID       261      304       By similarity.
FT DISULFID       461      496       By similarity.
FT CARBOHYD       218      218       N-linked (GlcNAc...) (Potential).
SQ SEQUENCE       498 AA; 54880 MW; 7658511BC01A8A01 CRC64;

Query Match          100.0%; Score 78; DB 1; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.00017; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0;

QY 1 AYHGYWQDDIYSL 13
DB 99 AYHGYWQDDIYSL 111

RESULT 2
Q76CT3 PRELIMINARY; PRT; 498 AA.
AC Q76CT3;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Alpha-amylase.
GN Name=amyA;
OS Aspergillus kawachi (Aspergillus awamori var. kawachi);
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=40384;
RN [1]
RP SEQUENCE FROM N.A.
RA Ito K.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB109452; BAD01053.1; -.
DR GO; GO:0004556; P:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PRO0110; ALPHAAMYLASE.
DR SMART; SM00642; Amy; 1.
SQ SEQUENCE 498 AA; 54679 MW; 8407BBFC7F4E3FC0 CRC64;

Query Match          100.0%; Score 78; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
DB 98 AYHGYWQDDIYSL 110

RESULT 3
AMYA ASPOR STANDARD; PRT; 499 AA.
AC P10529; P11763; Q00250;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA) (1,4-
DE Alpha-D-glucan glucanohydrolase).
GN Name=AMY1;
GN and
GN Name=AMY2;

```

```

GN and
GN Name=AMY3;
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 63303;
RX MEDLINE=89237897; PubMed=2785629;
RA Witsel S., Lachmund A., Wildhardt G., Rutkowski E.;
RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
RT intron-exon organization.";
RL Mol. Microbiol. 3:3-14(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89378767; PubMed=2789162; DOI=10.1016/0378-1119(89)90096-6;
RA Genes M.J., Dove M.J., Sellig V.L.;
RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
RT containing eight introns.";
RL Gene 79:107-117(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene of
RT Aspergillus oryzae.";
RL Agric. Biol. Chem. 53:593-599(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128276; PubMed=2612911; DOI=10.1016/0378-1119(89)90506-4;
RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A., Udaoka S.;
RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
RT evidence for multiple related genes.";
RL Gene 84:319-327(1989).
RN [5]
RP SEQUENCE OF 22-499.
RA Toda H., Kondo K., Narita K.;
RT "The complete amino acid sequence of Taka-amylase A.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
RN [6]
RP SEQUENCE OF 206-225.
RX MEDLINE=74001521; PubMed=4733850;
RA Isemura S., Ikenaka T.;
RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
RT A with trypsin and chymotrypsin.";
RL J. Biochem. 74:1-10(1973).
RN [7]
RP SEQUENCE OF 433-499.
RA Narita K.;
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=8027691; PubMed=6156152;
RA Matsuura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
RA Toda H., Narita K., Kakudo M.;
RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
RT 3-A resolution.";
RL J. Biochem. 87:1555-1558(1980).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=84212370; PubMed=6609921;
RA Matsuura Y., Kusunoki M., Harada W., Kakudo M.;
RT "Structure and possible catalytic residues of Taka-amylase A.";
RL J. Biochem. 95:697-702(1984).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
RX MEDLINE=97428212; PubMed=9283074; DOI=10.1021/bi970539i;
RA Brzozowski A.M., Davies G.J.;
RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
RT inhibitor acarbose at 2.0-A resolution.";
RL Biochemistry 36:10837-10845(1997).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.

```

CC -I- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations.
CC -I- SUBUNIT: Monomer.
CC -I- BIOTECHNOLOGY: Used in the brewing industry to increase the
CC fermentability of beer worts (including those made from unmalted
CC cereals), in the starch industry to make high maltose and high DE
CC syrups (starch saccharification), in the alcohol industry to
CC reduce fermentation time, in the cereal food industry for flour
CC supplementation and improvement of chilled and frozen dough, and
CC in the forestry industry for low-temperature modification of
CC starch. Sold under the name Fungamyl by Novozymes.
CC -I- MISCELLANEOUS: The sequence of Amy1 and Amy2 is shown.
CC -I- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X12725; CAA31218.1; -;
CC EMBL; X12726; CAA31219.1; -;
CC EMBL; X12727; CAA31220.1; -;
CC EMBL; D00434; BAA00336.1; -;
CC EMBL; M33218; AAA32708.1; -;
CC PIR; JK0201; JKO201.
CC PIR; JT0466; JTO466.
CC PIR; S04548; ALASI.
CC PDB; 2TRA; X-ray; A=22-499.
CC PDB; 6TRA; X-ray; @=22-499.
CC PDB; 7TRA; X-ray; @=22-499.
CC GlycoSuiteDB; P10529; -;
CC InterPro; IPR006589; Alp_amy1_cat_sub.
CC InterPro; IPR006047; Alpha_amy1_cat.
CC Pfam; PF00128; Alpha-amy1ase; 1.
CC PRINTS; PR00110; ALPHAAMYLAASE.
CC SMART; SM00642; Amy; 1.
KW 3D-structure: Calcium-binding; Carbohydrate metabolism;
KW Direct protein sequencing; Glycoprotein; Glycosidase; Hydrolase;
KW Multigene family; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499 Alpha-amy1ase A.
FT ACT_SITE 227 297 Nucleophile.
FT ACT_SITE 251 251 Proton donor.
FT ACT_SITE 318 318
FT METAL 142 142 Calcium 1.
FT METAL 183 183 Calcium 1 (via carbonyl oxygen).
FT METAL 196 196 Calcium 1.
FT METAL 227 227 Calcium 2.
FT METAL 231 231 Calcium 1 (via carbonyl oxygen).
FT METAL 251 251 Calcium 2.
FT DISULFID 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
FT
FT VARIANT 56 56 N-linked (GlcNAc...).
FT VARIANT 172 172 /FTID=CAR_000125.
FT CONFLICT 93 94 Q -> R (in AMY3).
FT CONFLICT 106 106 F -> L (in AMY3).
FT CONFLICT 184 184 TT -> DC (in Ref. 5).
FT CONFLICT 184 184 Q -> T (in Ref. 5).
FT CONFLICT 195 195 D -> Y (in Ref. 3).
FT CONFLICT 255 255 P -> L (in Ref. 3).
FT CONFLICT 291 291 G -> V (in Ref. 3).
FT CONFLICT 291 291 D -> H (in Ref. 4).
FT CONFLICT 345 345 I -> L (in Ref. 5).
FT CONFLICT 370 370 L -> A (in Ref. 4).
FT CONFLICT 406 409 WPIY -> PYI (in Ref. 5).
FT CONFLICT 448 448 G -> S (in Ref. 5).
FT CONFLICT 497 497 S -> SD (in Ref. 5 and 7).

FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 68
FT TURN 69 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 114
FT HELIX 118 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 148
FT HELIX 150 152
FT HELIX 155 157
FT STRAND 159 159
FT STRAND 164 166
FT HELIX 167 167
FT STRAND 172 172
FT TURN 176 177
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 192 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT HELIX 229 231
FT TURN 234 235
FT HELIX 236 244
FT TURN 245 245

Query Match 100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYYQQDIYSL 13
|||||
DB 99 AYHGYYQQDIYSL 111

RESULT 4
AMYB ASPAW
ID AMYB ASPAW STANDARD; PRT; 499 AA.
AC Q02906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amy1ase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE Glucanohydrolase B).
GN Name=AMYB;
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVKI43F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L., Kodama K.H.,

```

RA Royer T.J., Thompson S.A., Ward M., Wilson L.J., Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RL from Aspergillus niger var. awamori.";
CC Curr. Genet. 17:203-212(1990).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52756; CAA36967.1; -.
CC PIR; B48305; B48305.
CC
CC HSSP; P10529; 7TAA.
CC InterPro; IPR006589; Alp_amyl_cat_sub.
CC InterPro; IPR006047; Alpha_amyl_cat.
CC InterPro; IPR006046; Glyco_hydro_13.
CC Pfam; PF00128; Alpha-amylase; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Aamy; 1.
CC Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolyase; Multigene family; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227 Nucleophile (By similarity).
FT ACT_SITE 251 251 Proton donor (By similarity).
FT ACT_SITE 318 318 By similarity.
FT ACT_SITE 318 318 By similarity.
FT METAL 142 142 Calcium 1 (By similarity).
FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
FT METAL 183 183 similarity).
FT METAL 196 196 Calcium 1 (By similarity).
FT METAL 227 227 Calcium 2 (By similarity).
FT METAL 231 231 Calcium 1 (via carbonyl oxygen) (By
FT METAL 231 231 similarity).
FT METAL 251 251 Calcium 2 (By similarity).
FT DISULFID 51 59 By similarity.
FT DISULFID 171 185 By similarity.
FT DISULFID 304 304 By similarity.
FT DISULFID 461 496 By similarity.
FT CARBOHYD 218 218 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

Query Match 100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00017; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0;

QY 1 AYHGYYQQDIYSL 13
Db |||||
99 AYHGYYQQDIYSL 111

RESULT 5
ID AMY ASPSH STANDARD; PRT; 499 AA.
AC P30292;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN Name=AMY;
OS Aspergillus shirousami.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5070;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 92323146; PubMed-1368777;
RA Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT expression in Saccharomyces cerevisiae.";
RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10461; BAA01255.1; -.
CC HSSP; P10529; 7TAA.
CC InterPro; IPR006589; Alp_amyl_cat_sub.
CC InterPro; IPR006047; Alpha_amyl_cat.
CC InterPro; IPR006046; Glyco_hydro_13.
CC Pfam; PF00128; Alpha-amylase; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Aamy; 1.
CC Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolyase; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227 Nucleophile (By similarity).
FT ACT_SITE 251 251 Proton donor (By similarity).
FT ACT_SITE 318 318 By similarity.
FT ACT_SITE 318 318 By similarity.
FT METAL 142 142 Calcium 1 (By similarity).
FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
FT METAL 183 183 similarity).
FT METAL 196 196 Calcium 1 (By similarity).
FT METAL 227 227 Calcium 2 (By similarity).
FT METAL 231 231 Calcium 1 (via carbonyl oxygen) (By
FT METAL 231 231 similarity).
FT METAL 251 251 Calcium 2 (By similarity).
FT DISULFID 51 59 By similarity.
FT DISULFID 171 185 By similarity.
FT DISULFID 261 304 By similarity.
FT DISULFID 461 496 By similarity.
FT CARBOHYD 218 218 N-linked (GlcNAc...) (By similarity).
SQ SEQUENCE 499 AA; 54852 MW; 1FB7A850DA01C03F CRC64;

Query Match 100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00017; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0;

QY 1 AYHGYYQQDIYSL 13
Db |||||
99 AYHGYYQQDIYSL 111

RESULT 6
ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Taka-amylase A (EC 3.2.1.1).
DE Name=amyA;
GN Aspergillus oryzae.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]

```


RP SEQUENCE FROM N.A.
RC STRAIN=RIB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Goni K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amylolytic gene expression in *Aspergillus*
RT *oryzae*.";
RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB021876; BAA95703.1; -;
DR PIR; JN0588; JN0588.
DR PIR; S04549; ALAS3.
DR HSP; P10529; 7TAA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp amyl cat sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Amy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
DB |||||||||||
99 AYHGYWQDIYSL 111

RESULT 7
Q76L99 PRELIMINARY; PRT; 499 AA.
AC Q76L99;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amylase precursor.
GN Name=amy1;
OS *Aspergillus awamori*.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT-11;
RA Matsubara T.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB083159; BAD06002.1; -;
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp amyl cat sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Amy; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 499 AA; 54794 MW; 2B357AE3B836C3C2 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
DB |||||||||||
99 AYHGYWQDIYSL 111

us-10-820-200-2_copy_98_110.rup

Db 99 AYHGYWQDIYSL 111

RESULT 8
Q7LV45 PRELIMINARY; PRT; 499 AA.
AC Q7LV45;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amylase (EC 3.2.1.1).
GN Name=amy1;
OS *Aspergillus flavus*.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
OX NCBI_TaxID=5059;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86-10D;
RA Fakhoury A.M., Woloshuk C.P.;
RT "Amy1, the alpha-amylase gene of *Aspergillus flavus*: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914(1999).
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AF139925; AAF14264.1; -;
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp amyl cat sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Amy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
DB |||||||||||
99 AYHGYWQDIYSL 111

RESULT 9
Q9UV07 PRELIMINARY; PRT; 490 AA.
AC Q9UV07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 26, Last annotation update)
DE Alpha-amylase AmyA.
GN Name=amyA;
OS *Emicella nidulans* (*Aspergillus nidulans*).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; *Emicella*.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RA Boase N.A., Murphy R.M., Kelly J.M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AF208225; AAF17103.1; -;
DR HSP; P10529; 7TAA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp amyl cat sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.

us-10-820-200-2_copy_98_110.rup

```

DR SMART; SM00642; Aamy; 1.
SQ SEQUENCE 490 AA; 54249 MW; A891C4ACEABE5305 CRC64;
  Query Match 96.2%; Score 75; DB 2; Length 490;
  Best Local Similarity 92.3%; Pred. NO. 0.00053;
  Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
  |||||
Db 91 AYHGYWQDDIYAL 103

RESULT 10
ID 015751 PRELIMINARY; PRT; 128 AA.
AC 015751;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AmyA (Fragment).
GN Name=amyA;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RC SEQUENCE FROM N.A.
RA Loomis W.F., Iranfar N.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF020284; AAB70852.1; -.
DR HSSP; P10529; 7TAA.
DR DictyBase; DDB0214924; amyA.
DR GO; GO:0004556; F.alpha-amylose activity; IEA.
DR GO; GO:0005975; P.carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amy_cat.
DR Pfam; PF00128; Alpha-amyase; 1.
FT NON TER 1
FT NON TER 128
FT NON TER 128
SQ SEQUENCE 128 AA; 14950 MW; 2ACB8FDC55E79637 CRC64;

Query Match 87.2%; Score 68; DB 2; Length 128;
Best Local Similarity 83.3%; Pred. NO. 0.0018;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YHGYWQDDIYSL 13
  |||||
Db 3 YHGYWQDDIYTV 14

RESULT 11
AMY3 SCHPO STANDARD; PRT; 513 AA.
AC 014154; Q874R5; Q96WR6;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable alpha-amylose meu30 precursor (EC 3.2.1.1) (1,4-alpha-D-
DE glucan glucanohydrolase) (Meiotic expression up-regulated protein 30).
GN Name=mdes5; Synonyms=meu30; ORFNames=SPAC25H1.09, SPAC4A8.01;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., William R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.W.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

```

```

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jegels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Money P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Huer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armetrong J., Foreburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 191-513 FROM N.A.
RC STRAIN=CD16-1;
RX MEDLINE=21270454; PubMed=11376151; DOI=10.1093/nar/29.11.2327;
RA Watanabe T., Miyashita K., Saito T.T., Yoneki T., Kakiyama Y.,
RA Nabeshima K., Kishi Y.A., Shimoda C., Nojima H.;
RT "Comprehensive isolation of meiosis-specific genes identifies novel
RT proteins and unusual non-coding transcripts in Schizosaccharomyces
RT pombe."
RL Nucleic Acids Res. 29:2327-2337(2001).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; Z98762; CAB11471.1; -.
DR EMBL; Z98944; CAB62442.1; -.
DR EMBL; AB054314; BAB60880.1; -.
DR PIR; T38770; T38770.
DR HSSP; P10529; 7TAA.
DR GeneDB SPombe; SPAC25H1.09; -.
DR InterPro; IPR006047; Alpha_amy_cat.
DR InterPro; IPR006589; Alp_amy_cat_sub.
DR Pfam; PF00128; Alpha-amyase; 1.
DR SMART; SM00642; Aamy; 1.
DR Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolase; Meiosis; Signal.
FT SIGNAL 1 25
FT CHAIN 26 513
FT ACT SITE 226 226 Probable alpha-amyase meu30.
FT ACT SITE 226 226 Nucleophile (By similarity).
FT ACT SITE 250 250 Proton donor (By similarity).
FT ACT SITE 318 318 By similarity.
FT METAL 143 143 Calcium 1 (By similarity).
FT METAL 182 182 Calcium 1 (via carbonyl oxygen) (By
FT METAL 182 182 similarity).
FT METAL 195 195 Calcium 1 (By similarity).
FT METAL 226 226 Calcium 2 (By similarity).
FT METAL 230 230 Calcium 1 (via carbonyl oxygen) (By
FT METAL 230 230 similarity).
FT METAL 250 250 Calcium 2 (By similarity).
FT DISULFID 52 60 By similarity.
FT DISULFID 171 184 By similarity.
FT DISULFID 260 304 By similarity.

```

```

FT DISULFID 454 488 By similarity.
FT CARBOHYD 162 162 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 357 357 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 513 AA; 58715 MW; 455DD97FA428C182 CRC64;

Query Match 85.9%; Score 67; DB 1; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.012;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AHGYWQDDIYSL 13
Db 100 AHGYWQDDIYSL 112

RESULT 12
Q92394 PRELIMINARY; PRT; 631 AA.
AC Q92394;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-amylase precursor.
DE Names=amy-CS2;
GN Cryptococcus sp. S-2.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellomycetidae Incertae sedis; Cryptococcus.
OX NCBI_TaxID=87049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-2;
RX MEDLINE=96433120; PubMed=8836148;
RA Iefuji H., Chino M., Kato M., Iimura Y.;
RT "Raw-starch-digesting and thermostable alpha-amylase from the yeast
RT Cryptococcus sp. S-2: purification, characterization, cloning, and
RT sequencing.";
RL Biochem. J. 318:989-996(1996).
DR EMBL; D83541; BAA12011.1; -.
DR EMBL; D83540; BAA12010.1; -.
DR PIR; S72270; S72270.
DR HSSP; P10529; 7TAA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR Pfam; PF00128; Alpha-amylase; 1.
DR Pfam; PF00686; CBM 20; 1.
DR ProDom; PD001568; Glyco_hydro_CBD; 1.
DR SMART; SM00642; Amy; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 1 20 alpha-amylase.
SQ SEQUENCE 631 AA; 67658 MW; 8196B7B6E1D707E5 CRC64;

Query Match 84.6%; Score 66; DB 2; Length 631;
Best Local Similarity 76.9%; Pred. No. 0.022;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AHGYWQDDIYSL 13
Db 105 AHGYWQDDIYSL 117

RESULT 13
AMYA_ASPNG STANDARD; PRT; 484 AA.
AC P56271;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Acid alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;

```

```

RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=91002514; PubMed=2207069;
RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
RA Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.;
RT "Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-
RT A resolution of two enzymes from Aspergillus.";
RL Biochemistry 29:6244-6249(1990).
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -|- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations.
CC -|- SUBUNIT: Monomer.
CC -|- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
DR PDB; 2AAA; X-ray; @=1-484.
DR InterPro; IPR006589; Alp_aml_cat_sub.
DR InterPro; IPR006047; Alpha_aml_cat.
DR Pfam; PF00128; Alpha-amylase; 1.
DR SMART; SM00642; Amy; 1.
KW 3D-structure; Calcium-binding; Carbohydrate metabolism; Glycoprotein;
KW Glycosidase; Hydrolase.
FT ACT_SITE 206 206 Nucleophile.
FT ACT_SITE 230 230 Proton donor.
FT ACT_SITE 297 297
FT METAL 121 121 Calcium 1.
FT METAL 162 162 Calcium 1 (via carbonyl oxygen).
FT METAL 175 175 Calcium 1.
FT METAL 206 206 Calcium 2.
FT METAL 210 210 Calcium 1 (via carbonyl oxygen).
FT METAL 230 230 Calcium 2.
FT DISULFID 30 38
FT DISULFID 150 164
FT DISULFID 240 283
FT DISULFID 440 475
FT CARBOHYD 24 24 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 157 157 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 197 197 N-linked (GlcNAc... ) (Potential).
FT HELIX 3 6
FT TURN 7 8
FT STRAND 11 13
FT HELIX 16 19
FT TURN 22 23
FT HELIX 32 34
FT HELIX 42 47
FT TURN 48 48
FT HELIX 49 53
FT TURN 54 56
FT TURN 59 62
FT STRAND 66 68
FT STRAND 73 73
FT TURN 74 75
FT STRAND 76 76
FT TURN 79 80
FT STRAND 84 90
FT TURN 92 94
FT HELIX 97 108
FT TURN 109 111
FT STRAND 113 118
FT STRAND 122 122
FT STRAND 125 125
FT HELIX 129 131
FT STRAND 134 136
FT STRAND 138 138
FT HELIX 143 145
FT STRAND 146 146
FT STRAND 151 151
FT TURN 155 156
FT HELIX 158 163
FT STRAND 165 167
FT STRAND 172 173
FT STRAND 175 176
FT TURN 178 179
FT HELIX 181 198

```


RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KT-11;
 RA Matsubara T.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB083162; BAD06005.1; -;
 DR EMBL; AB083160; BAD06003.1; -;
 DR HSSP; P04064; IACZ.
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR006047; Alpha_amyl_cat.
 DR InterPro; IPR006589; Alp_amyl_cat_sub.
 DR InterPro; IPR002044; Glyco_hydro_CBD.
 DR Pfam; PF00128; Alpha-amylase; 1.
 DR Pfam; PF00686; CBM_20; 1.
 DR ProDom; PD001568; Glyco_hydro_CBD; 1.
 DR SMART; SM00642; Aamy; 1.
 KW Signal.
 FT SIGNAL 1 21 Potential.
 SQ SEQUENCE 634 AA; 69242 MW; 16C0B6AF6FB0E9B CRC64;

Query Match 82.1%; Score 64; DB 2; Length 634;
 Best Local Similarity 76.9%; Pred. No. 0.047;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDIYSL 13
 |||||
 Db 99 AYHGYWQKIYDV 111

Search completed: February 8, 2005, 18:41:53
 Job time : 8.62934 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2005, 14:22:56 ; Search time 7392 Seconds
(without alignments)
11366.524 Million cell updates/sec

Title: US-10-820-200-1
Perfect score: 1734
Sequence: 1 tcacatcaagctctcccttc.....aaaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	1734	100.0	AX140291 Sequence
2	1689.6	97.4	D10461 A. shirouea
3	1633.2	94.2	AB083159 Aspergill
4	1083	62.5	X12725 Aspergillus
5	1078.8	62.2	AB109452 Aspergill
6	1065.4	61.4	D00434 Aspergillus
7	1063.8	61.3	E02659 DNA sequenc
8	994	57.3	X12726 Aspergillus
9	992.4	57.2	X12727 Aspergillus
10	992.4	57.2	AB021876 Aspergill
11	989	57.0	M33218 A. oryzae Ta
12	985.8	56.9	AF139925 Aspergill
13	979	56.5	X52756 A.niger amy
14	908.2	52.4	X52755 A.niger amy
15	759.8	43.8	AX711606 Sequence
16	669.8	38.6	AB083160 Aspergill
17	660	38.1	E09025 cdna encodi
18	461.2	26.6	AF155463 Lipomyces
19	434.8	25.1	U30376 Lipomyces k

20	409	23.6	2143	8	SCSWA2	X73497 S.occidenta
21	377.4	21.8	4215	6	E01174	E01174 DNA encodin
22	377.4	21.8	4223	8	SFALP1	X05791 Saccharomyc
23	377	21.7	1404	6	E03536	E03536 DNA sequenc
24	375.4	21.6	1404	6	E08057	E08057 DNA encodin
25	375.4	21.6	1404	6	E08060	E08060 DNA encodin
26	375.4	21.6	1404	6	E23833	E23833 Sequence 7
27	375.4	21.6	1404	6	E23833	E23833 Sequence 4
28	375.4	21.6	1404	6	E08058	E08058 DNA encodin
29	373.8	21.6	1404	6	E08058	E08058 DNA encodin
30	373.8	21.6	1404	6	E08059	E08059 DNA encodin
31	373.8	21.6	1404	6	E23830	E23830 Sequence 1
32	373.8	21.6	1404	6	E23831	E23831 Sequence 3
33	373.8	21.6	1404	6	E23832	E23832 Sequence 5
34	373.8	21.6	1404	6	E23831	E23831 Sequence 1
35	373.8	21.6	1404	6	E23832	E23832 Sequence 2
36	373.8	21.6	1404	6	E23833	E23833 Sequence 3
37	373.8	21.6	1404	6	E23830	E23830 Sequence 1
38	373.8	21.6	1404	6	E23831	E23831 Sequence 2
39	373.8	21.6	1404	6	E23832	E23832 Sequence 3
40	371.4	21.4	2160	8	S77586	S77586 Debaryomyce
41	371.4	21.4	2161	8	SOAMY	X62079 S. occiden
42	369.2	21.3	2293	8	SOAMY1	X16040 Schwanniom
43	369.2	21.3	4190	6	AX591010	AX591010 Sequence
44	369.2	21.3	4190	6	AX591163	AX591163 Sequence
45	369.2	21.3	4190	6	AX717586	AX717586 Sequence

ALIGNMENTS

RESULT 1
LOCUS AX140291 1734 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 1 from Patent WO0134784.
ACCESSION AX140291
VERSION AX140291.1 GI:14280549
KEYWORDS Aspergillus oryzae
SOURCE Aspergillus oryzae
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1
AUTHORS bisg Rd-Frantzen,H., Svendsen,A. and Pedersen,S.
TITLE Fungamyl-like alpha-amylase variants
JOURNAL Patent: WO 0134784-A 1 17-MAY-2001;
Novozymes A/S (DK)
FEATURES
source
1..1734
/organism="Aspergillus oryzae"
/mol_type="unassigned DNA"
/db_xref="taxon:5062"
54..1550
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC40348.1"
/translation="MVAVMSFLYGLQVAAPALATPADWRSQSIYFLLTDRFARTDG
STTATCNTADOKYCGTQWGIIDKLDYIQMGFTAIWITPVAQLPQTAYGDAYHGY
WOODIYLNENYGTADDLKALSSALHERGMYLMDVVAHMGYDAGSSVDYSVFKPF
SSQDIFHPFCFIQNYEDQVEDCWLGNVTSLPDDDTDKDVVKNWYDVGSLVSNY
STDGURIDITKHVRQDFWPKYKAAGVYICIGVLDPDPAYTCPYQNVMDVGLNPIY
PLNNAFKSTSGSMDDLNTKSDPCDSTLLGTFFVENHDPNFRFASYNNDALAKNV
AAFIILNDGIPILYIAQEOHYAGGNDPANREATWLSGYPTDSELYKLIIASANAIRNVA
LSKDTGFVYTKNWPYIKDQTTIAMRKGTGDSQIIVTILSNKSGSDSYTILSLGAGTA
GQQLTEVIGCTTIVTGSNGNVPVPMAGGLPRVLYPTEKLAGSKICSSS"
114..1547
/product="unnamed"

mat_peptide

ORIGIN

Query Match 100.0%; Score 1734; DB 6; Length 1734;
Best Local Similarity 100.0%; Fred. No. 0;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TCACATCAAGCTCTCCCTTCTCTGAACATAAACCACAGAGGCAATTTATGATGGTGG 60
Db |||||
1 TCACATCAAGCTCTCCCTTCTCTGAACATAAACCACAGAGGCAATTTATGATGGTGG 60
QY 61 CTTGTGGTCTCTATTTCTGTATAGGCTTTCAGGTGCGGCACTGCTGTGGCTGCAAGCG 120
Db |||||
61 CTTGTGGTCTCTATTTCTGTATAGGCTTTCAGGTGCGGCACTGCTGTGGCTGCAAGCG 120
QY 121 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTACGGATCGATTGTCAGAGCGG 180
Db |||||
121 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTACGGATCGATTGTCAGAGCGG 180
QY 181 ATGGGTTCGACGACTGGGCTGTGTAATACCTGCGGATCAGAAATACCTGTGTGGAACATGGC 240
Db |||||
181 ATGGGTTCGACGACTGGGCTGTGTAATACCTGCGGATCAGAAATACCTGTGTGGAACATGGC 240
QY 241 AGGGCATCATGACAAAGTTGGAATATATCCAGGGAATGGCTTTCACAGCCATCTGGATCA 300
Db |||||
241 AGGGCATCATGACAAAGTTGGAATATATCCAGGGAATGGCTTTCACAGCCATCTGGATCA 300
QY 301 CCCCCTTACAGCCAGCTGCCCGACAGACCAACCGCATATGGAGATGCTTACCATGGCTACT 360
Db |||||
301 CCCCCTTACAGCCAGCTGCCCGACAGACCAACCGCATATGGAGATGCTTACCATGGCTACT 360
QY 361 GGCAGCAGGATATATATCTCTGAAACGAAACTACGGCACTGCAGATGACTTGAAGGCGC 420
Db |||||
361 GGCAGCAGGATATATATCTCTGAAACGAAACTACGGCACTGCAGATGACTTGAAGGCGC 420
QY 421 TCTTCTCGGCCCTTATGAGAGGGGATGTATCTTATGGTCGATGGTGTGTAACCATTA 480
Db |||||
421 TCTTCTCGGCCCTTATGAGAGGGGATGTATCTTATGGTCGATGGTGTGTAACCATTA 480
QY 481 TGGCTATGATGAGCGGCTAGCTCAGTCGATACAGTGTGTTAAACCGTTCAGTTCCC 540
Db |||||
481 TGGCTATGATGAGCGGCTAGCTCAGTCGATACAGTGTGTTAAACCGTTCAGTTCCC 540
QY 541 AAGACTACTTCCACCGGTTCTGTTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG 600
Db |||||
541 AAGACTACTTCCACCGGTTCTGTTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG 600
QY 601 ATTGCTGGCTAGGAGATAACACTGTCTCTTCCCTGCTGATCTCGATACCAACCAAGGATGG 660
Db |||||
601 ATTGCTGGCTAGGAGATAACACTGTCTCTTCCCTGCTGATCTCGATACCAACCAAGGATGG 660
QY 661 TCAGAAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTTGACGGCC 720
Db |||||
661 TCAGAAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTTGACGGCC 720
QY 721 TCCGATTCGACACAGTAAACACGCTCCAGAGGACTTCTGGCCGGGTACAAACAAAGCCG 780
Db |||||
721 TCCGATTCGACACAGTAAACACGCTCCAGAGGACTTCTGGCCGGGTACAAACAAAGCCG 780
QY 781 CAGGCGTGTACTGTATCGCGAGGCTCTCGACGGTGATCCGGCTACACTTGTCCCTACC 840
Db |||||
781 CAGGCGTGTACTGTATCGCGAGGCTCTCGACGGTGATCCGGCTACACTTGTCCCTACC 840
QY 841 AGAAGCTCATGACGCGGCTACTGAACATATCCCATTTTACTATCCACTCTCAACGCCCTTCA 900
Db |||||
841 AGAAGCTCATGACGCGGCTACTGAACATATCCCATTTTACTATCCACTCTCAACGCCCTTCA 900
QY 901 AGTCAACCTCCGCGACGATGGAGCTCTTACAACATGATCAACCGGTCAATCCGACT 960
Db |||||
901 AGTCAACCTCCGCGACGATGGAGCTCTTACAACATGATCAACCGGTCAATCCGACT 960
QY 961 GTCCAGACTCAACACTCTCGGCACATTCGTTCGAGAACCAACAGCAACCCACCGTTTCGCTT 1020
Db |||||
961 GTCCAGACTCAACACTCTCGGCACATTCGTTCGAGAACCAACAGCAACCCACCGTTTCGCTT 1020
QY 1021 CTTACACCAACGACATAGCCCTCGGCAAGAACGTCGAGGATTCATCTCCTCAACGACG 1080
Db |||||
1021 CTTACACCAACGACATAGCCCTCGGCAAGAACGTCGAGGATTCATCTCCTCAACGACG 1080
```

```
QY 1081 GAATCCCCCATCATCTACCCCGGCCAAGAACAGCACTACGCCGCGGAAACGACCCCGCA 1140
Db |||||
1081 GAATCCCCCATCATCTACCCCGGCCAAGAACAGCACTACGCCGCGGAAACGACCCCGCA 1140
QY 1141 ACCGGAAGCAACCTGGCTCTCGGCTACCCGACGACGAGCTGTACAGTTAAATTG 1200
Db |||||
1141 ACCGGAAGCAACCTGGCTCTCGGCTACCCGACGACGAGCTGTACAGTTAAATTG 1200
QY 1201 CTTCCGGAACCAATCCGGAATATGCAATAGCAAAAGATACAGATTTCGTGACCTACA 1260
Db |||||
1201 CTTCCGGAACCAATCCGGAATATGCAATAGCAAAAGATACAGATTTCGTGACCTACA 1260
QY 1261 AGAATCGCCCCATCTCAAAAGACGACACAAACGATCGCCATGCGCAAGGCAAGATGGGT 1320
Db |||||
1261 AGAATCGCCCCATCTCAAAAGACGACACAAACGATCGCCATGCGCAAGGCAAGATGGGT 1320
QY 1321 CCGAGATCGTACTATCTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCTCTCCT 1380
Db |||||
1321 CCGAGATCGTACTATCTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCTCTCCT 1380
QY 1381 TGAGTGGTGGGTTTACACAGCCGCCAGCAATTCACCGAGGTCAATTCGCTGCACGACCG 1440
Db |||||
1381 TGAGTGGTGGGTTTACACAGCCGCCAGCAATTCACCGAGGTCAATTCGCTGCACGACCG 1440
QY 1441 TGACGGTGGTTCGATGGAATGTCCCTTCTATGSCAGGTGGGCTACCTAGGGTAT 1500
Db |||||
1441 TGACGGTGGTTCGATGGAATGTCCCTTCTATGSCAGGTGGGCTACCTAGGGTAT 1500
QY 1501 TGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGTGGAGA 1560
Db |||||
1501 TGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGTGGAGA 1560
QY 1561 GTATATGATGGTACTGCTATTCATCTGSCATTTGACAGTGTGAGTTGATGTACA 1620
Db |||||
1561 GTATATGATGGTACTGCTATTCATCTGSCATTTGACAGTGTGAGTTGATGTACA 1620
QY 1621 GTTGAGTGGTACTGCTGCTATCCCTTATCTCTTCGATTGTTTTCGAACCCCTAAATG 1680
Db |||||
1621 GTTGAGTGGTACTGCTGCTATCCCTTATCTCTTCGATTGTTTTCGAACCCCTAAATG 1680
QY 1681 CCAAGCAGCTAGTCTATTTATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
Db |||||
1681 CCAAGCAGCTAGTCTATTTATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
```

RESULT 2

```
ASNAMY 1721 bp mRNA linear PLN 01-FEB-2000
LOCUS A. shirousamii mRNA for alpha-amylase, complete cds.
DEFINITION D10461 D11380
ACCESSION D10461.1 GI:217806
VERSION alpha-amylase; starch inducible.
KEYWORDS Aspergillus shirousamii
SOURCE Aspergillus shirousamii
ORGANISM Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 1721)
AUTHORS Shibuya,I., Tamura,G., Ishikawa,T. and Hara,S.
TITLE Cloning of the alpha-amylase cDNA of Aspergillus shirousamii and
its expression in Saccharomyces cerevisiae
JOURNAL Biosci. Biotechnol. Biochem. 56 (2), 174-179 (1992)
MEDLINE 92323146
PUBMED 1368777
REFERENCE 2 (bases 1 to 1721)
AUTHORS Shibuya,I.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1992) Ichiro Shibuya, Institute for Production
Research and Development, Nikka Whisky Distilling Co., Ltd.; 967
Azanatsuyama, Kaashiwa, Chiba 277, Japan (Tel:0471-72-3900,
Fax:0471-75-0290)
COMMENT Submitted (20-Jan-1992) to DDBJ by:
Ichiro Shibuya
Institute for Production
```


Research and Development
967 Azamatsu
Kaashiwa, Chiba 277
Japan
Phone: 0471-72-3900
Fax: 0471-75-0290.
Location/Qualifiers

FEATURES

source

1. .1721
/organism="Aspergillus shirousami"
/mol_type="mrna"
/strain="RIB 2504"
/db_xref="taxon:5070"

gene

1. .1721
/gene="AMY"

CDS

41..1540
/gene="AMY"
/EC number="3.2.1.1"
/codon_start=1
/product="alpha-amylase"
/protein_id="BA01255.1"
/db_xref="GI:217807"

/translations="MVAWMSFLYGLQVAAPALAAATPADMRSSQSIYFLLLDRFARD
GSTTATCNTADQKYCGGTWQGIIDKLDYIQGMGTAIWTPVTAQIPQTTAYGDAYHG
YMQDIYSUNENFTADDLKLALSSALHERGMVLMVDVAVNHMGYDAGSSVDYVFKP
FSSQDYFHPFCFIONVEDOTVEDWLGONTVSLPDLDTTKDVVKNEWDMVGLSVN
YSIDGLIDTQVHKVQDFWPKAGVYICIGELVDVDPAYTCPYQNVMDGVLYPIY
YPLNLFKSTSGSMDLYNINVTVKSDFDSTLLGTFVENHDPFASFTINDLAKN
VAAFIILNDGIPILYAGSOHYAGNDPANREATWLSGYPTSELYKLJASANAIRNY
AISKDFVTYKMPYIKDITTIAMRKDGSIQIVTILSNKGASGDSYTLISLSGAGYT
AGQLTEVIGCTTVTGSDGNVPPMAGGLPLRVLYPTEKLAGSKICSSS"

sig_peptide

41..103
/gene="AMY"

mat_peptide

104..1537
/gene="AMY"
/product="alpha-amylase"
1721
/gene="AMY"

polya_site

ORIGIN

Query Match 97.4%; Score 1689.6; DB 8; Length 1721;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1698; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 11 CTCTCCCTCTCTGAACAATAAACCCACAGAGAGCATTTATGATGGTCGGTGGTGC 70
DB 1 CTCTCCCTCTCTGAACAATAAACCCACAGAGAGCATTTATGATGGTCGGTGGTGC 60
QY 71 TCTATTCTCTACGGCTTCAGGTTCGGCGCACCTGCTTTGGCTGCAACGCTCGGACTG 130
DB 61 TCTATTCTGTACGGCTTCAGGTTCGGCGCACCTGCTTTGGCTGCAACGCTCGGACTG 120
QY 131 GCGATCGCAATCCATTATTTCTCTCACGGATCGATTTGCAAGGACGGATGGGTCGAC 190
DB 121 GCGATCGCAATCCATTATTTCTCTCACGGATCGATTTGCAAGGACGGATGGGTCGAC 180
QY 191 GACTGGACTGTGTAATCTCGGATCAGAAATACGTGTGTGGAACATGGCAGGCAATCAT 250
DB 181 GACTGGACTGTGTAATCTCGGATCAGAAATACGTGTGTGGAACATGGCAGGCAATCAT 240
QY 251 CGACAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCCATCTGGATCACCCCGCTTAC 310
DB 241 CGACAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCCATCTGGATCACCCCGCTTAC 300
QY 311 AGCCAGCTGCCCCAGACACCGCATATGGAGATGCTTACCATGGCTACTGGCAGCAGGA 370
DB 301 AGCCAGCTGCCCCAGACACCGCATATGGAGATGCTTACCATGGCTACTGGCAGCAGGA 360
QY 371 TATATACTCTCTGAAAGAAATACGGCACTGAGATGACTTGAAGGCGCTCTCTTCGGC 430
DB 361 TATATACTCTCTGAAAGAAATACGGCACTGAGATGACTTGAAGGCGCTCTCTTCGGC 420
QY 431 CCTTCATGAGAGGGGATGATCTTATGTCGATGGTGTCTTAAACCATATGGGCTATGA 490

DB 421 CCTTCATGAGAGGGGATGTATCTTATGTCGATGTGGTTGCTAACCATATGGGCTATGA 480
QY 491 TGGAGCGGCTAGCTCAGTCGATTTACAGTGTGTTTAAACCGTTCAGTTCCTCCAACTACTT 550
DB 481 TGGAGCGGCTAGCTCAGTCGATTTACAGTGTGTTTAAACCGTTCAGTTCCTCCAACTACTT 540
QY 551 CCACCCGTTCTGTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGGATTTGCTGGCT 610
DB 541 CCACCCGTTCTGTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGGATTTGCTGGCT 600
QY 611 AGGAGATAACACTGTCTCTCTGCTGATCTCGATACCAAGGATGTGTTCAAGATGA 670
DB 601 AGGAGATAACACTGTCTCTCTGCTGATCTCGATACCAAGGATGTGTTCAAGATGA 660
QY 671 ATGGTACGACTGGTGGGATCATTTGATTCGAACTACTCGATTCAGGCTTCGATTCGA 730
DB 661 ATGGTACGACTGGTGGGATCATTTGATTCGAACTACTCGATTCAGGCTTCGATTCGA 720
QY 731 CACAGTAAACACGCTCCAGAGACTTCTGGCCCGGGTACAAACAAAGCCGACGCGTGA 790
DB 721 CACAGTAAACACGCTCCAGAGACTTCTGGCCCGGGTACAAACAAAGCCGACGCGTGA 780
QY 791 CTGTATCGGCGAGGTCTCGAAGGCTGATCCGGCTACACTTGTCTTACAGAACGTCAT 850
DB 781 CTGTATCGGCGAGGTCTCGAAGGCTGATCCGGCTACACTTGTCTTACAGAACGTCAT 840
QY 851 GAGCGGCTACTGAACTATCCATTTACTATCCACTCTCAAGGCTTCAAGTCAACCTC 910
DB 841 GAGCGGCTACTGAACTATCCATTTACTATCCACTCTCAAGGCTTCAAGTCAACCTC 900
QY 911 CGGAGCATGGAGGACTCTTACAACTGATCAACCGTCAACCGTCAACCGTCAACCGTCA 970
DB 901 CGGAGCATGGAGGACTCTTACAACTGATCAACCGTCAACCGTCAACCGTCAACCGTCA 960
QY 971 AACACTCTCTGGGCACTTCTGAGAAACCAACCAACCGTTCGCTTCTTACACCAA 1030
DB 961 AACACTCTCTGGGCACTTCTGAGAAACCAACCAACCGTTCGCTTCTTACACCAA 1020
QY 1031 CGACATAGCCCTCGCCAAAGACGTCGACGATTCATCTCTCAACGACGAAATCCCAT 1090
DB 1021 CGACATAGCCCTCGCCAAAGACGTCGACGATTCATCTCTCAACGACGAAATCCCAT 1080
QY 1091 CATCTACGCGGCCAGAACGACACTACCGCGGGAAGACGACCCCGGAAACCGCGAAGC 1150
DB 1081 CATCTACGCGGCCAGAACGACACTACCGCGGGAAGACGACCCCGGAAACCGCGAAGC 1140
QY 1151 AACCTGGCTCTCGGGCTACCCGACGACGAGCTGTAAAGTAAATTCCTTCGCGAA 1210
DB 1141 AACCTGGCTCTCGGGCTACCCGACGACGAGCTGTAAAGTAAATTCCTTCGCGAA 1200
QY 1211 CGCAATTCGGAACATATGCCATTAGCAAGATACAGGATTCGTGACCTACAAGAACTGGCC 1370
DB 1201 CGCAATTCGGAACATATGCCATTAGCAAGATACAGGATTCGTGACCTACAAGAACTGGCC 1360
QY 1271 CATCTACAAGACGACAAACGATCGCCATGCGCAAGGACAGATGGGTTCGAGATCGT 1330
DB 1261 CATCTACAAGACGACAAACGATCGCCATGCGCAAGGACAGATGGGTTCGAGATCGT 1320
QY 1331 GACTATCTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCTCTCTCTTGAAGTGTGC 1390
DB 1321 GACTATCTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCTCTCTCTTGAAGTGTGC 1380
QY 1391 GGGTTACACAGCCGGCAGCAATTTGACGAGGTCAATTTGCTGCAACCGTGAACGTTGG 1450
DB 1381 GGGTTACACAGCCGGCAGCAATTTGACGAGGTCAATTTGCTGCAACCGTGAACGTTGG 1440
QY 1451 TTCGATGAAATGTCCTTCTTATGCGAGGTGGGCTACCTAGGGTATTTGATCCGAC 1510
DB 1441 TTCGATGAAATGTCCTTCTTATGCGAGGTGGGCTACCTAGGGTATTTGATCCGAC 1500
QY 1511 TCAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGTCTGTGAAGGTTGAGAGTATATGATG 1570
DB 1501 TGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGTCTGTGAAGGTTGAGAGTATATGATG 1560

Db 1043 TGTCTCTTTTACATCAATGAATGAATCTTGATGGTTTGTAGATACTCTCTGAACGAAAA 1102
|||
Qy 392 CTACGGCACTGACGATGACTTGAAGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTA 451
|||
Db 1103 CTACGGCACTGACGATGACTTGAAGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTA 1162
|||
Qy 452 TCTTATGGTCGATGTGGTTGCTTAACCATAT----- 481
|||
Db 1163 TCTTATGGTCGATGTGGTTGCTTAACCATATGGTTCGTGCTCTTTGCAACTGACTTCGGG 1222
|||
Qy 482 -----GGGCTATGATGGAGCGGGTAGC 503
|||
Db 1223 GATATGGTTCAUATTCAGTACTGACAATGAGTAATATCAGGGCTATGATGGAGCGGGTAGC 1282
|||
Qy 504 TCAGTCGATTACAGTGTGTTTAAACGGTTCAGTCCCAAGACTACTTCCACCCGTTCTGT 563
|||
Db 1283 TCAGTCGATTACAGTGTGTTTAAACGGTTCAGTCCCAAGACTACTTCCACCCGTTCTGT 1342
|||
Qy 564 TTCAATTCAAAACCTATGGAAGATCAGACTCAGGTTGAGGATTCGTGGCTAGGAGATAACACT 623
|||
Db 1343 TTCAATTCAAAACCTATGGAAGATCAGACTCAGGTTGAGGATTCGTGGCTAGGAGATAACACT 1402
|||
Qy 624 GTCTCCTTCCCTGATCTCGATACCAACAGGATGGTCAAGAAATGAATGGTACGACTGG 683
|||
Db 1403 GTCTCCTTCCCTGATCTCGATACCAACAGGATGGTCAAGAAATGAATGGTACGACTGG 1462
|||
Qy 684 GTGGGATCAATTCGTATCGAACTACTCCA----- 711
|||
Db 1463 GTGGGATCAATTCGTATCGAACTACTCCAAGTAAGATATTTCTCCCTCAATTCATAACTTGG 1522
|||
Qy 712 -----TTCACGGCTCCGCTATCGACACAGTAAACACAGT 745
|||
Db 1523 CTGATCGATGATCTTACGAATATGATTCAGCGCTCCGCTATCGACACAGTAAACACAGT 1582
|||
Qy 746 CCAGAAGGACTTCTGGCCCGGGTACAAACAAAGCCGAGCGGTGTACTGTATCGGCGAGGT 805
|||
Db 1583 CCAGAAGGACTTCTGGCCCGGGTACAAACAAAGCCGAGCGGTGTACTGTATCGGCGAGGT 1642
|||
Qy 806 GCTCGACGGTATCGCGCTACACTTGTCCCTACCGAAGCGTATGAGCGGGTACTGAA 865
|||
Db 1643 GCTCGACGGTATCGCGCTACACTTGTCCCTACCGAAGCGTATGAGCGGGTACTGAA 1702
|||
Qy 866 CTATCCCAT----- 874
|||
Db 1703 CTATCCCATGATGGTTCCTCCACCATGAGCTTCTTCCAAAGTCTCATCTCTTAACGAA 1762
|||
Qy 875 -----TTACTATCCACTCTCTCAAGCCCTTCAAGTCAACCTCCGGCGAGCATG 920
|||
Db 1763 ACGGCTAAAAACAGTTTACTATCTCACTCTCTCAAGCCCTTCAAGTCAACCTCCGGCGAGCATG 1822
|||
Qy 921 GAGGACCTCTACACATGATCAACACCGTCAATCCGACTGTCCAGACTCAACACTCCTG 980
|||
Db 1823 GAGGACCTCTACACATGATCAACACCGTCAATCCGACTGTCCAGACTCAACACTCCTG 1882
|||
Qy 981 GGCACATTCTGAGAACACGACACACACCGTTCGCTTC----- 1021
|||
Db 1883 GGCACATTCTGAGAACACGACACACCGTTCGCTTCGTTAAGTCTTCCTTTTATT 1942
|||
Qy 1022 -----TTACACCAACGACA 1035
|||
Db 1943 TTCCGTTCCCAATTTCCACACAGAACCCACCTTAACAGAGCAAGTTTACACCAAGACA 2002
|||
Qy 1036 TAGCCCTCCGCAAGAACGTCGAGGATTCATCTCTCAAGCGGAATCCCGATCATCT 1095
|||
Db 2003 TAGCCCTCCGCAAGAACGTCGAGGATTCATCTCTCAAGCGGAATCCCGATCATCT 2062
|||
Qy 1096 ACGCGGGCCAAAGAACGACTAGCGCGGGGAAACGACCCCGGACACCGGAGCAACCT 1155
|||
Db 2063 ACGCGGGCCAAAGAACGACTAGCGCGGGGAAACGACCCCGGAGACCGGAGCAACCT 2122
|||
Qy 1156 GGCTCTCGGGCTACCCGACCGACGAGCTGTATCAAGTTAATTTGCTCCGGGAACGCAA 1215
|||

RESULT 6
ASNTAA
LOCUS ASNTAA 2337 bp DNA linear PLN 29-MAY-2002
DEFINITION Aspergillus oryzae gene for Taka-amylase A precursor, complete cds.
ACCESSION D00434
VERSION D00434.1 GI:217822
KEYWORDS Taka-amylase A; alpha-1,4-glucan-4-glucanohydrolase.
SOURCE Aspergillus oryzae
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 2337)
AUTHORS Tada,S., Iimura,Y., Gomi,K., Takahashi,K., Hara,S. and Yoshizawa,K.
TITLE Cloning and nucleotide sequence of the genomic Taka-amylase A gene
of Aspergillus oryzae
JOURNAL Agric. Biol. Chem. 53, 593-599 (1989)
COMMENT Compared with the amino acid sequence of mature TAA reported by
Toda et al., Trp-385 is inserted and Asp-476 is deleted.
FEATURES
Location/Qualifiers
1..2337
/organism="Aspergillus oryzae"
/mol_type="genomic DNA"
/db_xref="taxon:5062"
/clone="phi-AA107"
/note="252 bp upstream of SalI site.-RIB40"
18..24
/note="putative"
TATA_signal
CDS
join(118..285,341..379,465..580,658..766,835..1063,
1122..1284,1350..1496,1562..1802,1882..2169)
/EC_number="3.2.1.1"

Db 788 ACTTCGGGATATGGTTCAATTTTCAGTACTGACAAATGAGTAATATCAGGCGCTATGATGGAG 847
QY 496 CGGGTAGCTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTTCCCAAGACTACTTCCACC 555
Db 848 CGGGTAGCTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTTCCCAAGACTACTTCCACC 907
QY 556 CGTTCTGTTTCAATTCAAAATATGAGATCAGACTCAGGTTGAGGATTCCTGGCTAGGAG 615
Db 908 CGTTCTGTTTCAATTCAAAATATGAGATCAGACTCAGGTTGAGGATTCCTGGCTAGGAG 967
QY 616 ATAACTGTCCTTCCTGCTGATCTCGATACACACAGGATGTGTCAGAAATGATGT 675
Db 968 ATAACTGTCCTTCCTGCTGATCTCGATACACACAGGATGTGTCAGAAATGATGT 1027
QY 676 ACAGCTGGGTGGGATCATTTGGTATCGAACTACTCCA----- 711
Db 1028 ACAGCTGGGTGGGATCATTTGGTATCGAACTACTCCAAGTAAGATATTTCTCCCTCATTTCTA 1087
QY 712 -----TTGACGGGCTCCGTTATCGACACAGTA 737
Db 1088 CAACCTTGGCTGATCGATGATCTTACGAAATCAGTTGACGGCTCCGTTATCGACACAGTA 1147
QY 738 AAACAGTCCAGAGACTTTCGGCCGGGTACAAAGCCGCGAGCGGTGTAATCTGTATC 797
Db 1148 AAACAGTCCAGAGACTTTCGGCCGGGTACAAAGCCGCGAGCGGTGTAATCTGTATC 1207
QY 798 GCGAGGTGCTCGACGGTATCGGCTACACTTGTCCCTACAGAACGTCATGACGCG 857
Db 1208 GCGAGGTGCTCGACGGTATCGGCTACACTTGTCCCTACAGAACGTCATGACGCG 1267
QY 858 GTACTGAACTATCCCAT----- 874
Db 1268 GTACTGAACTATCCCATGATGTTCTCCCAACCATGAGCCTTCTTGCAGTCTCATCTC 1327
QY 875 -----TTACTATCCATCTCTCAACGCTTCAAGCTCAACCTCG 912
Db 1328 CTAAAGAAACGGCTAAACAGTTACTATCCATCTCTCAACGCTTCAAGCTCAACCTCG 1387
QY 913 GCAGCATGACGACCTCTCAACATGATCAACACGTCGTAATCCGACTTCCAGACTCAA 972
Db 1388 GCAGCATGACGACCTCTCAACATGATCAACACGTCGTAATCCGACTTCCAGACTCAA 1447
QY 973 CACTCTCGGCGACATTCGTCGAGAACCGACCAACCCAGGTTGCTTC----- 1021
Db 1448 CACTCTCGGCGACATTCGTCGAGAACCGACCAACCCAGGTTGCTTCGTAAGTCTTCC 1507
QY 1022 -----TTACAC 1027
Db 1508 CTTTTTATTTTCGGTTCCCAATTTCCACAGAACCCCACTAAACAGAGCAAAAGTTTACAC 1567
QY 1028 CAACGACATAGCCTCGCCCAAGAACGTCGACGATTCATCATCTCAACGCGGAATCCC 1087
Db 1568 CAACGACATAGCCTCGCCCAAGAACGTCGACGATTCATCATCTCAACGCGGAATCCC 1627
QY 1088 CATCATCTACGCGGCGCAAGACAGCACTACGCGCGGGAACGACCCCGCGAACCGCGA 1147
Db 1628 CATCATCTACGCGGCGCAAGACAGCACTACGCGCGGGAACGACCCCGCGAACCGCGA 1687
QY 1148 AGCAACTGCTCTCGGCTACCCGACGACGAGCTGTCAAGTTAAATTCCTCGCG 1207
Db 1688 AGCAACTGCTCTCGGCTACCCGACGACGAGCTGTCAAGTTAAATTCCTCGCG 1747
QY 1208 GAACGCAATCCGGAATGATGCAATAGCAAGATACAGGATTCGTCACCTACA----- 1260
Db 1748 GAACGCAATCCGGAATGATGCAATAGCAAGATACAGGATTCGTCACCTACAAGTAAG 1807
QY 1261 ----- 1260
Db 1808 CACAACCTCTAAGCATACCTAATGGCTTATCTTCAGAGTATCTGACACAAGAGACTAAT 1867
QY 1261 -----AGAACTGCGCCCATCTTCAAGACGACACAAGATCGCCATCGCGAAG 1308
Db 1868 CACTGGCAATACAGAACTGCGCCCATCTCAAGAACGACACAAGATCGCCATCGCGAAG 1927

QY 1309 GCACAGATGGGTGCGAGATCGTGACTATCTTGTCCAAAGGGTCTTCGGGTGATTCGT 1368
Db 1928 GCACAGATGGGTGCGAGATCGTGACTATCTTGTCCAAAGGGTCTTCGGGTGATTCGT 1987
QY 1369 ATACCCTCTCCTTGTAGTGGTGGGTTACACAGCGCGCAGCAATGACGAGGTCATTG 1428
Db 1988 ATACCCTCTCCTTGTAGTGGTGGGTTACACAGCGCGCAGCAATGACGAGGTCATTG 2047
QY 1429 GCTGCACACCGTGACGGTTCGGATGGAATGTCCTTCTATGGCAGGTGGGC 1488
Db 2048 GCTGCACACCGTGACGGTTCGGATGGAATGTCCTTCTATGGCAGGTGGGC 2107
QY 1489 TACCTAGGGTATTTGATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTCTGT 1548
Db 2108 TACCTAGGGTATTTGATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTCTGT 2167
QY 1549 GAAGGTCGAGAGTATATGATGGTACTGCTATTTCAATCTGGCATTGGACAGTGGATTGA 1608
Db 2168 GAAGGTCGAGAGTATATGATGGTACTGCTATTTCAATCTGGCATTGGACAGTGGATTGA 2227
QY 1609 GTTTGATCTACAGTTGGAGTGGTACTGCTGTCATCCCTTATATCTTTCGATTGTTTTT 1668
Db 2228 GTTTGATCTACAGTTGGAGTGGTACTGCTGTCATCCCTTATATCTTTCGATTGTTTTT 2287
QY 1669 CGAACCTTAATGCCAAGCACGCTAGTCTATTATATAGGAAA 1707
Db 2288 CGAACCTTAATGCCAAGCACGCTAGTCTATTATATAGGAAA 2326
RESULT 7
E02659
LOCUS E02659 2337 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of takaamylase A gene.
ACCESSION E02659
VERSION E02659.1 GI:2170887
KEYWORDS JP 1990268685-A/1.
SOURCE Aspergillus oryzae
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 2337)
AUTHORS Takahashi,K., Iimura,M., Gomi,K., Hara,M., Yoshizawa,K., Tada,S.
and Tamura,G.
TITLE NOVEL GENE AND VECTOR, TRANSFORMING SUBSTANCE USING SAME GENE AND
VECTOR AND USE THEREOF
JOURNAL Patent: JP 1990268685-A 1 02-NOV-1990;
TAX ADM AGENCY, JOZO SHIGEN KENKYUSHO:KK
COMMENT OS Aspergillus oryzae
PN JP 1990268685-A/1
PD 02-NOV-1990
PF 07-APR-1989 JP 1989086787
PI TAKAHASHI KOJIRO, IIMURA MINORU, GOMI KATSUYA, HARA MASAMICHI,
YOSHIZAWA KIYOSHI, TADA SETSUZO, TAMURA GAKUZO PC
C12N15/56,C12G3/02,C12G3/02,C12N1/15,C12N9/30,C12N15/56, PC
C12R1:69).
PC (C12N1/15,C12R1:69), (C12N9/30,C12R1:69);
CC strandedness: single;
CC topology: linear;
CC hypothetical: No;
CC *source: strain=RIB40;
FH key Location/Qualifiers
FT CDS 1<..<2337
FT /product='takaamylase A'.
FEATURES
source
1..2337
/organism='Aspergillus oryzae'
/mol_type='genomic DNA'
/db_xref='taxon:5062'
ORIGIN
Query Match 61.3%; Score 1063.8; DB 6; Length 2337;

Best Local Similarity 75.3%; Pred. No. 2e-246; Matches 1700; Conservative 0; Mismatches 7; Indels 552; Gaps 8;			
QY	1	TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAAGGCAATTTATGATGGTCG	60
Db	68	TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAAGGCAATTTATGATGGTCG	127
QY	61	CGTGGTGGTCTCTATTTCTGTACGGCTTCAGGTCGGCGCACCTGTTGGCTGCAACGC	120
Db	128	CGTGGTGGTCTCTATTTCTGTACGGCTTCAGGTCGGCGCACCTGTTGGCTGCAACGC	187
QY	121	CTGCGACTGGCGATCGCAATCATTTATTTCTCTCAAGGATCGATTTGCAAGACGG	180
Db	188	CTGCGACTGGCGATCGCAATCATTTATTTCTCTCAAGGATCGATTTGCAAGACGG	247
QY	181	ATGGGTGCAAGCTGCGACTGTAATACTGCGGATC-----	216
Db	248	ATGGGTGCAAGCTGCGACTGTAATACTGCGGATCAGGTGTGTTGTACTACTAGCTT	307
QY	217	-----AGAAATACGTGTTGGTGGACATGGCAGGCG	245
Db	308	TCAGAAAGAGGAATGTAAACTGACTGATATAGAAATACGTGTTGGTGGACATGGCAGGCG	367
QY	246	ATCATCGACA-----	255
Db	368	ATCATCGACAAGGTAATTGCCCTTTATCAAAAAAAGAGGAAGACAGAGAGAAAAA	427
QY	256	-----AGTTGGACTATATCCAGGGAATGGG	280
Db	428	TAAATATAAAGAACTCTAGTCTCAACCATCACATAGTTGGACTATATCCAGGGAATGGG	487
QY	281	CTTCAGAGCAATCTGATCACCCCGCTTACAGCCGAGCTGCCACAGACACCGCATATGG	340
Db	488	CTTCAGAGCAATCTGATCACCCCGCTTACAGCCGAGCTGCCACAGACACCGCATATGG	547
QY	341	AGATGCTTACCATGGCTACTGGCAGCAGGATAT-----	373
Db	548	AGATGCTTACCATGGCTACTGGCAGCAGGATATGTAACGATCGTATTTCTTTGGAAATA	607
QY	374	-----ATACTCTCTG	383
Db	608	TCTACCTGTCTATTTCTACATCATATGAACATACTTGTATGTTTATAGATACTCTCTG	667
QY	384	AACGAAACTACGGCACTGAGATGATGAGGCGCTCTCTTCGGCCCTCTATGAGAGG	443
Db	668	AACGAAACTACGGCACTGAGATGATGATGAGGCGCTCTCTTCGGCCCTCTATGAGAGG	727
QY	444	GGGATGTATCTTATGTCGATGGTGTGCTAACCATAT-----	481
Db	728	GGGATGTATCTTATGTCGATGGTGTGCTAACCATATGGTTCGGTCCCTTTGCAACTG	787
QY	482	-----GGGCTATGATGGAG	495
Db	788	ACTTCGCGGATATGGTTTCACTTTCAGTACTGACAACTAGTAATATCAGGCTATGATGGAG	847
QY	496	CGGGTAGCTCAGTCGATTCAGTGTGTTAAACCGTTTCAGTTCGCCCTCTATGAGAGG	555
Db	848	CGGGTAGCTCAGTCGATTCAGTGTGTTAAACCGTTTCAGTTCGCCCTCTATGAGAGG	907
QY	556	CGTCTGTTTCAATCAAACTATGAGATCAGACTCAGGTTGAGGATTCCTGGCTAGGAG	615
Db	908	CGTCTGTTTCAATCAAACTATGAGATCAGACTCAGGTTGAGTATTTGCTGGCTAGGAG	967
QY	616	ATAAACACTGTCTCTTCCCTGATCTCGATACACCAAGGATGGTCAAGAAATGAATGGT	675
Db	968	ATAAACACTGTCTCTTCCCTGATCTCGATACACCAAGGATGGTCAAGAAATGAATGGT	1027
QY	676	ACGACTGGTGGGATCAATTTGGTATCGAACTATCCCA-----	711
Db	1028	ACGACTGGTGGGATCAATTTGGTATCGAACTATCCCAAGTAAGATATTTCTCCCTCATTTA	1087
QY	712	-----TTGACGGCTCCGTATCGACACAGTA	737
1088	CAACTTGGCTGATCGATGATACTTACGAAATCAGTTTGACGGCTCCGTATCGACACAGTA	1147	
738	AAACAGCTCCAGAAGACTTCTGGCCCGGGTACAAACAAAGCCGACGCGTGTACTGTATC	797	
1148	AAACAGCTCCAGAAGACTTCTGGCCCGGGTACAAACAAAGCCGACGCGTGTACTGTATC	1207	
798	GGCGAGGTCTCGACGGTGTATCCGGCTACACTTGTCCCTACCAGAACGTCATGGAGCGC	857	
1208	GGCGAGGTCTCGACGGTGTATCCGGCTACACTTGTCCCTACCAGAACGTCATGGAGCGC	1267	
858	GTACTGAACTATCCCAT-----	874	
1268	GTACTGAACTATCCCATGTATGGTCTCCAAACCATGAGCCTTCTTGCAAGTCTCATCTC	1327	
875	-----TTACTATCCACTCTCTCAACGCTTCAAGTCAACCTCCG	912	
1328	CTAAAGAAACGGCTAAACACAGTTACTATCCACTCTCTCAACGCTTCAAGTCAACCTCCG	1387	
913	GCAGCATGGACGACCTCTACAAACATGATCAACACCGTCAAAATCCGACTGTCCAGACTCAA	972	
1388	GCAGCATGGACGACCTCTACAAACATGATCAACACCGTCAAAATCCGACTGTCCAGACTCAA	1447	
973	CACCTCTGGGACATTCGTTCGAGAACCCAGCAACCCACGGTTCGCTTC-----	1021	
1448	CACCTCTGGGACATTCGTTCGAGAACCCAGCAACCCACGGTTCGCTTCGTAAGTCTTCC	1507	
1022	-----TTACAC	1027	
1508	CTTTTATTTTCGTTTCCCAATTTCCACACAGAACCCACCTAACAAAGAGCAAAAGTTACAC	1567	
1028	CAACGACATAGCCCTCGCCAGAAAGTGCAGCAGTTCATCATCTCAACAGACGGAATCCC	1087	
1568	CAACGACATAGCCCTCGCCAGAAAGTGCAGCAGTTCATCATCTCAACAGACGGAATCCC	1627	
1088	CATCATCTACGCGCGCCAAAGAACAGCACTACGCGCGGAAACGACCCCGGAAACCGCGA	1147	
1628	CATCATCTACGCGCGCCAAAGAACAGCACTACGCGCGGAAACGACCCCGGAAACCGCGA	1687	
1148	AGCAACTGGCTCTCGGGCTACCCGACCGACAGCAGTGTACAAATTAATTCCTCCGC	1207	
1688	AGCAACTGGCTCTCGGGCTACCCGACCGACAGCAGTGTACAAATTAATTCCTCCGC	1747	
1208	GAACGCAATCCGGAATATGCCATTTAGCAAAAGATACAGATTCGTGACCTACA-----	1260	
1748	GAACGCAATCCGGAATATGCCATTTAGCAAAAGATACAGATTCGTGACCTACAAGGTAAG	1807	
1261	-----	1260	
1808	CACAACCTCTAAGCATACCTAATGGCCTATCTTCAAGATATCTGACACAGAGACTAAT	1867	
1261	-----AGAACTGGCCCATCTACAAAGACGACACAAACGATTCGCCATGCGCAAGG	1308	
1868	CACCTGCAATACAGAACTGGCCCATCTACAAAGACGACACAAACGATTCGCCATGCGCAAGG	1927	
1309	GCACAGATGGGTGCGAGATCGTACTATCTTGTCCAAAGGGTGTTCGGGTGATTCGT	1368	
1928	GCACAGATGGGTGCGAGATCGTACTATCTTGTCCAAAGGGTGTTCGGGTGATTCGT	1987	
1369	ATACCTCTCTCTTGTAGTGGTGGGTTACACAGCGCGCCAGCAATTTGACGGAGGTCATTG	1428	
1988	ATACCTCTCTCTTGTAGTGGTGGGTTACACAGCGCGCCAGCAATTTGACGGAGGTCATTG	2047	
1429	GCTGACGACCGTGGTGGTTCGGATGGAAATGTGCTGTTCTTCTATGGCAGGTGGGC	1488	
2048	GCTGACGACCGTGGTGGTTCGGATGGAAATGTGCTGTTCTTCTATGGCAGGTGGGC	2107	
1489	TACCTAGGATATGATTCGCACTGAGAGTTGGCAGGTAGCAAGATCTCTGTAGTCTCGT	1548	
2108	TACCTAGGATATGATTCGCACTGAGAGTTGGCAGGTAGCAAGATCTCTGTAGTCTCGT	2167	
1549	GAAGGTGGAGAGTATATGATGCTATCTTCAATCTGGCATTCGACAGTGGATTCGA	1608	
2168	GAAGGTGGAGAGTATATGATGCTATCTTCAATCTGGCATTCGACAGTGGATTCGA	2227	

sig_peptide

QY 217 -----AGAAATACGTGTGGTGAACATGGCAGGC 245
DB TCAGAAAGAGGAATGTAAACTGACTTGATATAGAAATACGTGTGGTGAACATGGCAGGC 864
QY 246 ATCATCGACA----- 255
DB ATCATCGACAGGTAAATTGCCCTTTATCAAAAAAAGAGAAAGCAAGAAAGAA 924
QY 256 -----AGTTGGACTATATCCAGGGAATGG 280
DB TAAATAAAGAAAGAACTTAGTCTTAACCATACATAGTTGGACTATATCCAGGGAATGG 984
QY 281 CTTACAGCCATCTGATCACCCTGGTTACAGCCGAGTCCGCCAGCACCACCATATGG 340
DB CTTACAGCCATCTGATCACCCTGGTTACAGCCGAGTCCGCCAGCACCACCATATGG 1044
QY 341 AGATGCTTACCATGGTCTAGTGGCAGCAGGATAT----- 373
DB AGATGCTTACCATGGTCTAGTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC 1104
QY 374 -----ATACTCTCTGAACGAAA 391
DB TGTCACTTTTACATCAATATGAACATACTTGATGTTTTAGATACTCTCTGAACGAAA 1164
QY 392 CTAAGGCACTGAGATGACTTGAAGGGCTCTCTCGGCCCTTCATGAGAGGGGATGTA 451
DB CTAAGGCACTGAGATGACTTGAAGGGCTCTCTCGGCCCTTCATGAGAGGGGATGTA 1224
QY 452 TCTTATGCTGATGCTGCTAACCATAT----- 481
DB TCTTATGCTGATGCTGCTAACCATATGTTGCTGCTTTGCAACTGACTTCGGG 1284
QY 482 -----GGGCTATGATGAGCGGTAGC 503
DB GATATGTTTCACTTTCAGTACATGAGTAATATCAGGGCTATGATGGAGCGGTAGC 1344
QY 504 TCAGTCGATTAAGTGTGTTAAACCGTTCAAGTTCCTCAAGACTATCTTCAACCGCTCTGT 563
DB TCAGTCGATTAAGTGTGTTAAACCGTTCAAGTTCCTCAAGACTATCTTCAACCGCTCTGT 1404
QY 564 TTCATTCAAACTATCAAGATCAGACTCAGGTTGAGGATTCGCTGAGGATGAGGATACACT 623
DB TTCATTCAAACTATCAAGATCAGACTCAGGTTGAGGATTCGCTGAGGATGAGGATACACT 1464
QY 624 GTCTCCTTGCCTGATCTCGATACCAAGGATGCTGCTCAAGAAATGAATGGTACGACTGG 683
DB GTCTCCTTGCCTGATCTCGATACCAAGGATGCTGCTCAAGAAATGAATGGTACGACTGG 1524
QY 684 GTGGGATCATTTGGTATCGAACTACTCCA----- 711
DB GTGGGATCATTTGGTATCGAACTACTCCAAGTAAATATTTCTCCCTCATTTACAACCTGG 1584
QY 712 -----TTGACGCTCCGTATCGACACAGTAACACCT 745
DB CTGATCGATGATCTTACGAAATCAGTTGACGCTCCGTATCGACACAGTAACACCT 1644
QY 746 CCAGAAAGGACTTCTGGCCCGGTGACAAAGAGCCGAGCGGTGTACTGTATCGCGAGGT 805
DB CCAGAAAGGACTTCTGGCCCGGTGACAAAGAGCCGAGCGGTGTACTGTATCGCGAGGT 1704
QY 806 GTTCGACGGTGTATCCGGCTTACACTTGTCCCTACCAAGACGTCATCGACGGCGTACTGAA 865
DB GCTCGACGGTGTATCCGGCTTACACTTGTCCCTACCAAGACGTCATCGACGGCGTACTGAA 1764
QY 866 CTATCCCAT----- 874
DB CTATCCCATGATGTTCTCCCAACCATGAGCCTTCTTCAAGTCTCATCTCTCAACGAA 1824
QY 875 -----TTACTATCCACTCTCAAGCCTTCAAGTCAACCTCCGGCAGCATG 920
DB ACGGCTAAACCGATTACTATCACTCTCAAGCCTTCAAGTCAACCTCCGGCAGCATG 1884

QY 921 GACGACTCTACACATGATCAACACCGTCAAAATCCGACTGTCCAGACTCAACACTCCTG 980
DB GACGACTCTACACATGATCAACACCGTCAAAATCCGACTGTCCAGACTCAACACTCCTG 1944
QY 981 GGCACATTCGTTCGAGAAACACGACCAACCCACGGTTCGCTTC----- 1021
DB GGCACATTCGTTCGAGAAACACGACCAACCCACGGTTCGCTTCGTAAGTCTTCCTTTTATT 2004
QY 1022 -----TTACACCAAGACA 1035
DB TTCCGTTTCCCAATTTCCACACAGAACCCCACTTAACAGAGCAAAAGTTACACCAAGACA 2064
QY 1036 TAGCCCTTCGCAAGACGTCGAGCATTCATCTCTCAACGACGAAATCCCATCATCT 1095
DB TAGCCCTTCGCAAGACGTCGAGCATTCATCTCTCAACGACGAAATCCCATCATCT 2124
QY 1096 AGCCGGCCCAAGAACAGCACTACGCGCGGGAACGACCCCGGAACCGGGAAGCAACCT 1155
DB AGCCGGCCCAAGAACAGCACTACGCGCGGGAACGACCCCGGAACCGGGAAGCAACCT 2184
QY 1156 GGTCTTCGGGTACCCGACCGACAGCGAGCTGTACAAAGTTAAATTGGCTTCGCGAAGCA 1215
DB GGTCTTCGGGTACCCGACCGACAGCGAGCTGTACAAAGTTAAATTGGCTTCGCGAAGCA 2244
QY 1216 TCGGAACATATGCCATTAGCAAGATACAGGATTCGTGACCTACA----- 1260
DB TCGGAACATATGCCATTAGCAAGATACAGGATTCGTGACCTACA----- 2304
QY 1261 ----- 1260
DB CTAAGCATACCTTAATGGCTATCTTCAGAGTATCTGCACAAAGAGACTAATCACTGGCA 2364
QY 1261 -----AGAACTGGCCCATCTTACAAAGACGACACAAAGATCGCCATGCGAAGGCA 1316
DB ATACAGAACTGGCCCATCTTACAAAGACGACACAAAGATCGCCATGCGAAGGCA 2424
QY 1317 GGTTCGAGATCGTGAATCTTGTTCACAAAGGGTCTTCGGGTGATTCGTATACCTC 1376
DB GGTTCGAGATCGTGAATCTTGTTCACAAAGGGTCTTCGGGTGATTCGTATACCTC 2484
QY 1377 TCTTGTAGTGTGGGTTTACACAGCGCGCCAGCAATTCACGAGGTCTATGGCTGCACG 1436
DB TCTTGTAGTGTGGGTTTACACAGCGCGCCAGCAATTCACGAGGTCTATGGCTGCACG 2544
QY 1437 ACCGTGACGGTTGGTTTCGGATGGAATGTCCTGTTCTTATGCGAGTGGGCTACCTAG 1496
DB ACCGTGACGGTTGGTTTCGGATGGAATGTCCTGTTCTTATGCGAGTGGGCTACCTAG 2604
QY 1497 GTATTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTACGTCGTAAGGGTG 1556
DB GTATTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTACGTCGTAAGGGTG 2664
QY 1557 GAGATGATATGATGTTACTGCTATTCAATCTGCAATTCGACAGTGGTGGTGGTGGTG 1616
DB GAGATGATATGATGTTACTGCTATTCAATCTGCAATTCGACAGTGGTGGTGGTGGTG 2724
QY 1617 TA 1618
DB 2725 TA 2726

RESULT 9

AOAMY3
LOCUS Aspergillus oryzae amy3 gene for alpha-amylase (EC 3.2.1.1).
DEFINITION
ACCESSION X12727.1 GI:2454
VERSION X12727.1
KEYWORDS amy3 gene; amylase-alpha.
SOURCE Aspergillus oryzae
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE
1 (bases 1 to 3146)

Db 1164 CTACGGCACTGCAGATGACTTGAAGGGCTCTCTTCGGCCCTTCATGAGAGGGGATGTA 1223
QY 452 TCTTATGGTCGATGTGGTTGCTTAACCATAT- 481
Db 1224 TCTTATGGTCGATGTGGTTGCTTAACCATATATGGTTCTGGTTCCTTTGCAACTGACTTTCGG 1283
QY 482 -----GGGCTATGATGGAGCGGTAGC 503
Db 1284 GATATGGTTCAATTCAGTACTGCAATGAGTAATATCAGGGCTATGATGGAGCGGTAGC 1343
QY 504 TCAGTCGATTACAGTGTGTTAAACCGTTCCAGTTCCTCAAGACTACTTCCACCCGTTCTGT 563
Db 1344 TCAGTCGATTACAGTGTGTTAAACCGTTCCAGTTCCTCAAGACTACTTCCACCCGTTCTGT 1403
QY 564 TTCATTCAAACTATGAAGATCAGACTCAGGTTGAGGATTCCTGGCTAGGAGATAACACT 623
Db 1404 CTCATTCAAACTATGAAGATCAGACTCAGGTTGAGGATTCCTGGCTAGGAGATAACACT 1463
QY 624 GTCTCCTTGCCTGATCTCGATACCAACCAAGGATGTGTCAGAAATGAATGGTTACGACTGG 683
Db 1464 GTCTCCTTGCCTGATCTCGATACCAACCAAGGATGTGTCAGAAATGAATGGTTACGACTGG 1523
QY 684 GTGGGATCATTTGGTATCGAACTACTCCA----- 711
Db 1524 GTGGGATCATTTGGTATCGAACTACTCCAAGTAAGATATTTCTCCCTCATTTCAAACTGG 1583
QY 712 -----TTGACGGCCTCCGCTATCGACACAGTAAACACAGT 745
Db 1584 CTGATCGATGATATTTACGAAATCAGTTGACGGCCTCCGTTATCGACACAGTAAACACAGT 1643
QY 746 CCAGAAAGGACTTCTGCGCCGGGTACAAACAAAGCCGAGCGGTGTACTGTATCGCGAGGT 805
Db 1644 CCAGAAAGGACTTCTGCGCCGGGTACAAACAAAGCCGAGCGGTGTACTGTATCGCGAGGT 1703
QY 806 GCTCGACGGTGATCCGCGCTACACTTGTCCCTACCAAGAGCTCATGACGGCGTACTGAA 865
Db 1704 GCTCGACGGTGATCCGCGCTACACTTGTCCCTACCAAGAGCTCATGACGGCGTACTGAA 1763
QY 866 CTATCCCAT----- 874
Db 1764 CTATCCCATGTATGGTTCTCCCAACCATGAGCCTCTTTCGAAGTCTCATCTCTTAAGAA 1823
QY 875 -----TTACTATCCACTCTCAAGCCTTCAAGTCAACCTCCCGCAGCATG 920
Db 1824 ACGGCTAAACACCACTTACTATCCACTCTCAAGCCTTCAAGTCAACCTCCCGCAGCATG 1883
QY 921 GACGACTCTTACACATGATCAACCGTCAAAATCCGACTGTCCGACTCAACACTCCCTG 980
Db 1884 GACGACTCTTACACATGATCAACCGTCAAAATCCGACTGTCCGACTCAACACTCCCTG 1943
QY 981 GGCACATTCGTTCGAGAACCAACCAACCCACCGTTGCTTC----- 1021
Db 1944 GGCACATTCGTTCGAGAACCAACCAACCCACCGTTGCTTCGCTTAAGTCTTCCCTTTTATT 2003
QY 1022 -----TTACACCAACGACA 1035
Db 2004 TTCGTTTCCCAATTTCCACACAGAACCCCACTTAACAGAGCAAAAGTTACACCAAGACA 2063
QY 1036 TAGCCCTCCGAGAACGTCGAGCATTCATCTCTCAACGCGAATCCCATCATCT 1095
Db 2064 TAGCCCTCCGAGAACGTCGAGCATTCATCTCTCAACGCGAATCCCATCATCT 2123
QY 1096 ACGCCGGCAAGAACGACTAGCCGGGGGAAACGACCCCGGAAACCGGAGCAACCT 1155
Db 2124 ACGCCGGCAAGAACGACTAGCCGGGGGAAACGACCCCGGAAACCGGAGCAACCT 2183
QY 1156 GGCTCTCGGCTACCCGACGACGAGCTGTACAAAGTTAAATGGCTCCGGAACGCAA 1215
Db 2184 GGCTCTCGGCTACCCGACGACGAGCTGTACAAAGTTAAATGGCTCCGGAACGCAA 2243
QY 1216 TCCGGAACCTATGCCATTAGCAAGATACAGGATTCGTGACCTACA----- 1260
Db 2244 TCCGGAACCTATGCCATTAGCAAGATACAGGATTCGTGACCTACAAGGTAAGCAACACT 2303

QY 1261 ----- 1260
Db 2304 CTAAGCATACCCCTAATGGCTATCTTCAAGTATCTGACACAGAGACTAATCACTGGCA 2363
QY 1261 ----AGAACTGGCCCATCTACAAAGACGACACAAACGATCGCCATGCGCAAGGCGACAGAT 1316
Db 2364 ATACAGAACTGGCCCATCTACAAAGACGACACAAACGATCGCCATGCGCAAGGCGACAGAT 2423
QY 1317 GGGTCGCAGATCGTGACTACTTCTGTCCAAACAAAGGGTCTTCGGGTGATTCTGTTATACCCCTC 1376
Db 2424 GGGTCGCAGATCGTGACTACTTCTGTCCAAACAAAGGGTCTTCGGGTGATTCTGTTATACCCCTC 2483
QY 1377 TCTTGAGTGGTGGGTTTACACAGCCGCCAGCAATTTGACGAGGTCAATTGGCTGTCACG 1436
Db 2484 TCTTGAGTGGTGGGTTTACACAGCCGCCAGCAATTTGACGAGGTCAATTGGCTGTCACG 2543
QY 1437 ACCGTGACGGTTGGTTCGGATGGAAATGTCCTTCTTATGCGAGTGGCTACCTAGG 1496
Db 2544 ACCGTGACGGTTGGTTCGGATGGAAATGTCCTTCTTATGCGAGTGGCTACCTAGG 2603
QY 1497 GTATTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 1556
Db 2604 GTATTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 2663
QY 1557 GAGAGTATATGATGTTGCTGCTTATTCATCTGGCATTTGACAGTGGAGTTGAGTTTGATG 1616
Db 2664 GAGAGTATATGATGTTGCTGCTTATTCATCTGGCATTTGACAGTGGAGTTGAGTTTGATG 2723
QY 1617 TA 1618
Db 2724 TA 2725

RESULT 10

AB021876
LOCUS
DEFINITION
AB021876 35244 bp DNA linear PLN 17-MAY-2000
Aspergillus oryzae amyR, agda, amyA gene for amyRp (transcriptional activator), alpha-glucosidase, Taka-amylase A, complete cds.

AB021876
VERSION
KEYWORDS
SOURCE
ORGANISM
AB021876.1 GI:7939499
amyR; alpha-glucosidase; Taka-amylase A.
Aspergillus oryzae
Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE
AUTHORS
TITLE
1 (sites)
Gomi,K., Akeno,T., Minetoki,T., Ozeki,K., Kumagai,C., Okazaki,N.
and Iimura,Y.
Molecular cloning and characterization of a transcriptional
activator gene, amyR, involved in the amyolytic gene expression in
Aspergillus oryzae

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
Aspergillus oryzae
BioSci Biotechnol. Biochem. 64 (4), 816-827 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
Submitted (25-DEC-1998) Kateuya Gomi, Tohoku University, Graduate
School of Agricultural Science; 1-1, Tsutsumidori-Amamiyanachi,
Aoba-ku, Sendai, Miyagi 981-8555, Japan
(E-mail:gomi@biochem.tohoku.ac.jp, Tel:81-22-717-8901,
Fax:81-22-717-8902)

FEATURES
source
1. 35244
/organism="Aspergillus oryzae"
/mol_type="genomic DNA"
/strain="RIB40"
/db_xref="taxon:5062"
complement(join(11172..11549,11634..12084,12151..13136))
/gene="amyR"
complement(join(11172..11549,11634..12084,12151..13136))
/gene="amyR"

```
/function="transcriptional activator"
/codon_start=1
/product="amyRp"
/protein_id="BAA95701.1"
/db_xref="GI:7939500"
/translation="MSHSPDIPSTSEKEMESTPBKPKQACDNCRRRIKCSRELPC
KTCQRLLSQSVSDVLRKPKFRLYPLAHPILASRPLRPLTKWLPNPGACHLAS
PTSPSTVADADQLHPDFSESPRLPPDLVSSPDSTNSLFDSSITGAPAPRLSTP
NLLAHNVPLKYLFIMPVVRQDQOQCHQPERLSPQRYAFIAALCAATHQLKLDG
AAPGPEASARASLDGHPMLSGEELLAEARARKEYNVVDEINMELLTSLFLFAVG
NLDROQAWFYLCQTSWVFTGLQRETSYKLSVEBAEERVRFWLLPVTVERGVALQ
QAPVMLRNSIHKPOLVGCDDPILAYGINLININVEKLSPNLDWVSAGSSADGPP
PTSSQSLAKQISLEGVSEIOKVDILITQWLMKLSMTHVTQCGSRDDAYLFP
HLPLVGLKAVMGVITLRSRGOSYLPALVEOSRGIISFDCSLISDPLPFGGPPAI
MWRTEGSPDLLGIADDLQERENEGGEGIVAGEEISF"
gene
Join(14807..15507,15567..15977,16030..16200,16259..17933)
Join(14807..15507,15567..15977,16030..16200,16259..17933)
CDS
/gene="agda"
Join(14807..15507,15567..15977,16030..16200,16259..17933)
/gene="agda"
/codon_start=1
/product="alpha-glucosidase"
/protein_id="BAA95702.1"
/db_xref="GI:7939501"
/translation="WAGLKSFLLASSWLLPVACGASQSIIVPSTSAAYSQFTIPASD
VGANIVANIDDPQAVNAQSVCPGYKASDVKSSQGFSTASLELAGDPQNVYGTVDVSLT
LTVEYQADRNLQIIVPTVFDASNASWVILSEILVPRPKASQNASVPQSDFFVWSNE
PSFNEFKRKATGDVLTNKGSTLVENQFIEFVTLPEENLYGLGERMQLRLLEN
ANLTIYADIADIPDNNYGHAFYLDTRYKYKVGQNSHTIIVKSSBAEPQSEYVYS
HGVFURNAGHQEILLRDLQIWRITLGGSDVLTFFSYGTOAEVTKQYLSLTVGLPAMQY
YNTLGHQCRWAGNNSPFEVDIANFERFEIPELVMADIDYMHGYRNFNDQHRFSY
BEGEFKLNKLHAGNSRVPVIVDGLYIIPNENASDAYETDRGAKDVFVIFKNDGSLY
IGMWPVTVYDPMHPRKASDPWANELVTWNKLHYDGVYDMAVSVFCVGSCTGPN
LSMNPAHPPLPFGPQNVVDYPSGFNITNATKASASAGASOASASSTTTAPY
LATPTFGVRNVDPYVINYVQPHDLSVHAI SPNTHSDGVQYDVHSLYGHGIN
IYHGLLKWENKRPFIITARLSTFGSGKWAGHGDGNSFKMGSMFFSISQALQSLFG
ITPMGVDCTGNGNDLNCNRMQLSAFPPEYRNHNVLSAIPQEPYMAVSDATKA
AMNIYAILPYETFLPHLHTTGTVMRALAWEFFNDPSLAAGVQFVLVPSVMVIVP
LEPOVDYVQGVPGVGHVYDYSOTAVDAKPGVNTTISAPLGHIPVEYVRGGSILP
MQEVALTFRDARKTPWLSLASSNGTASGQLYLDGSEVYPEDTSLVDPLASRSLR
ASARGTKEANPLANVTYLVGTETKPSSTVTLNGETLSSDSVKYNATSHVLHVGLQKHT
ADGAWAKDWLKW"
repeat_unit
18036..19912
/note="transposon-like element"
gene
Join(20440..20607,20663..20701,20787..20902,20972..21080,
21149..21377,21436..21598,21664..21810,21876..22116,
22196..22483)
/gene="amyA"
Join(20440..20607,20663..20701,20787..20902,20972..21080,
21149..21377,21436..21598,21664..21810,21876..22116,
22196..22483)
CDS
/gene="amyA"
/EC_number="3.2.1.1"
/codon_start=1
/product="Taka-amyIase A"
/protein_id="BAA95703.1"
/db_xref="GI:7939502"
/translation="NMVAWSLFLYGLQVAPALAAATPADWFSQSIYFLITDRPARD
GSTATCNTADRYKCGGTQWQIILKLDIYQMGFTAIWITPVTAOLPQTAYGDYHG
WYQDIYSLNENYATDADLKALSSALHERGYMLVMDVNVANHWGSDGSSVDYVFKP
FSSQDYFHPFCLIQNYEQTVEDCWLGDNTVSLPDLTDTKDVVKNYMDWVSIVSN
YSIDGLRDTVKHOKDFWPGYKAAGVYICIGVLDGDPAYTCPYQNYMDVGLNPIY
VPLLAFKSTSGSDDLNMINTVNKVSDCPDSTLLGTFFVENHDPNPFASVTNDIALKN
VAAFIILNDGIPILYAGQEQHVAGGNDPANREATWLSGVPTDSELYKLIASANARNY
ALSKDTGFVYKWNPIYKDDTTIAMRKDGTGSQIVTILSNKASGSDSTLSLGSAGYT
AGQQUTEVIGYKTTVTYVGSNDVFPVMAGGLPRVLYPTEKLAGSKTICSSS"
ORIGIN
Query Match 57.2%; Score 992.4; DB 8; Length 35244;
Best Local Similarity 74.8%; Pred. No. 5.9e-229;
Matches 1617; Conservative 0; Mismatches 1; Indels 544; Gaps 8;
.QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCACTTATGATGGTCG 60
```

```
Db 21470 CCAGAGGACTTCTGCGCCGGGTACAAACGCGGAGCGTGTACTGTATCGCGGAGGT 21529
QY 806 GCTCGACGGTGATCCGGCTTACACTTGTCCCTACCAAGACGTATGACGCGGTACTGAA 865
Db 21530 GCTCGACGGTGATCCGGCTTACACTTGTCCCTACCAAGACGTATGACGCGGTACTGAA 21589
QY 866 CTATCCCAT----- 874
Db 21590 CTATCCCATGTATGGTTCCTCCAAACGATGAGCCCTTCTGCAAGTCTCATCTCCTAACGAA 21649
QY 875 -----TTACTATCACTCCTCAAGCCCTTCAAGTCAACCTCCGCGAGCATG 920
Db 21650 ACGGCTAAACCAAGTTACTATCACTCCTCAAGCCCTTCAAGTCAACCTCCGCGAGCATG 21709
QY 921 GACGACCTCTAACAATGATCAACACGCTCAAAATCGAGCTGTCCAGACTCAACACCTCTG 980
Db 21710 GACGACCTCTAACAATGATCAACACGCTCAAAATCGAGCTGTCCAGACTCAACACCTCTG 21769
QY 981 GGCACATTCTGTCGAGAACCAACCAACCGCTTCGCTTC----- 1021
Db 21770 GGCACATTCTGTCGAGAACCAACCAACCGCTTCGCTTC-----TTACACCAACGACA 21829
QY 1022 -----TTACACCAACGACA 1035
Db 21830 TTCGTTGCCAATTTCCACACGAACCCCACTTAAACGAGCAAGTTACACCAACGACA 21889
QY 1036 TAGCCCTCGCAAGAACGTCGAGCATTCATCTCTCAACGACGAATCCCATCATCT 1095
Db 21890 TAGCCCTCGCAAGAACGTCGAGCATTCATCTCTCAACGACGAATCCCATCATCT 21949
QY 1096 ACGCCGGCCAAAGACGACTACGCGCGGGAACGACCCCGGACCGGAGCAACCT 1155
Db 21950 ACGCCGGCCAAAGACGACTACGCGCGGGAACGACCCCGGACCGGAGCAACCT 22009
QY 1156 GGCCTCGGGCTACCGGACGAGCGAGCTGTACAAAGTTAAATGCTCGCGAAGCGCA 1215
Db 22010 GGCCTCGGGCTACCGGACGAGCGAGCTGTACAAAGTTAAATGCTCGCGAAGCGCA 22069
QY 1216 TCCGAACTATGCCATTAGCAAGATACAGGATTCGTGACCTACA----- 1260
Db 22070 TCCGAACTATGCCATTAGCAAGATACAGGATTCGTGACCTACAAGGTAAGCAACCT 22129
QY 1261 ----- 1260
Db 22130 CTAAGCATACCTTAATGGCCTATCTTTCAGAGTATCTGACAAAGACTAATCACTGGCA 22189
QY 1261 ----AGAACTGGCCCATCTACAAAGACGACACAAACGATCGCAAGGGGACAGAT 1316
Db 22190 ATACAGAACTGGCCCATCTACAAAGACGACACAAACGATCGCAAGGGGACAGAT 22249
QY 1317 GGGTCGAGATCGTACTATCTTGTCCAAAGGGTGCTTCGGGGTGATTCGTATACCCCTC 1376
Db 22250 GGGTCGAGATCGTACTATCTTGTCCAAAGGGTGCTTCGGGGTGATTCGTATACCCCTC 22309
QY 1377 TCGTTGAGTGGTCGGGTTACAGCGCGGACCAATTTGACGAGGTCAATGGCTGACG 1436
Db 22310 TCGTTGAGTGGTCGGGTTACAGCGCGGACCAATTTGACGAGGTCAATGGCTGACG 22369
QY 1437 ACCGTGACGTTGGTTCGGATGGAATGTCCTGTTCCATGTCAGTGGCTGACCTAG 1496
Db 22370 ACCGTGACGTTGGTTCGGATGGAATGTCCTGTTCCATGTCAGTGGCTGACCTAG 22429
QY 1497 GTATTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 1556
Db 22430 GTATTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 22489
QY 1557 GAGAGTATATGAGTGTACTGCTATTCAATCTGGCAATGGACAGTGTGAGTTGATG 1616
Db 22490 GAGAGTATATGAGTGTACTGCTATTCAATCTGGCAATGGACAGTGTGAGTTGATG 22549
QY 1617 TA 1618
Db 22550 TA 22551
```

RESULT 11

```
ASNTAAG1
LOCUS       A.oryzae Taka-amylase A (Taa-GI) gene, complete cds.
DEFINITION  A.oryzae Taka-amylase A (Taa-GI) gene, complete cds.
ACCESSION  M33218.1
VERSION    M33218.1 GI:166530
SOURCE     Taka-amylase A.
ORGANISM   Aspergillus oryzae
            Aspergillus oryzae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE  1 (bases 1 to 2935)
AUTHORS    Tsukagoshi,N., Furukawa,M., Nagaba,H., Kiritu,N., Tauboi,A. and
            Uda,K.
TITLE      Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
            evidence for multiple related genes
JOURNAL    Gene 84 (2), 319-327 (1989)
MEDLINE    90128276
PUBMED     2612911
COMMENT    Original
FEATURES   Location/Qualifiers
            source
            1..2935
            /organism="Aspergillus oryzae"
            /mol_type="genomic DNA"
            /db_xref="taxon:5062"
            prim_transcript
            543..>2789
            /note="Taa-GI mRNA and introns"
            join(607..772,828..868,955..1070,1140..1248,1317..1545,
            1603..1765,1830..1976,2041..2281,2360..2644)
            /note="Taka-amylase A (Taa-GI) precursor"
            /codon_start=1
            /protein_id="AAA32708.1"
            /db_xref="GI:166531"
            /translation="MMVAWMSIFLYCLOVAPALATPADWRSQSIYFLLTDRPARD
            GSTTATCNADQKCYCGTQWQIILDKLDYIQHGFTAIWTPVTAQLPQTAYGDAYHG
            YWQDIYSLENIYQADDDLKUSSALHGMVLMVDVAVNHMGYDGAGSSVDYSVFKP
            FSSQYFHPFLIQNYEDQVDEWLGNTVSLPDLTDTKDYVKNKWDVWGSLSVN
            YSIDGLRIDTVKHQDFWPGYKAGVYICIGVLDGDPAYTCFQVNVMDGVNLYPIY
            YLLNAFKSTSGMDLNMINTKVDSPDCPDLLGTFEVNHDPNRPASVYNDLALAKN
            VAAFIILANDGIPILYAGQEHYAGNDPANRENTWASGYTDSLEYKLILASNAIRNY
            ALSQDTFVYKNWFIYKDDTTIARKKGTGDSQIIVTLNKGASGSYTLSLSGAGYT
            ACQQTETVIGTCTTIVGSDGNVFPVMAGGLRVLYPTEKLAGSKICSSS"
            sig_peptide
            /note="Taka-amylase A signal peptide"
            join(670..772,828..829)
            /product="Taka-amylase A"
            mat_peptide
            join(772,828..868)
            /product="Taka-amylase A"
            mat_peptide
            join(955..1070,1140)
            /product="Taka-amylase A"
            mat_peptide
            join(1069..1070,1140..1248)
            /product="Taka-amylase A"
            mat_peptide
            join(1317..1545,1603..1604)
            /product="Taka-amylase A"
            mat_peptide
            join(1545,1603..1765,1830)
            /product="Taka-amylase A"
            mat_peptide
            join(1764..1765,1830..1976,2041)
            /product="Taka-amylase A"
            mat_peptide
            join(1975..1976,2041..2281)
            /product="Taka-amylase A"
            mat_peptide
            2360..2644
            /product="Taka-amylase A"
            exon
            607..772
            /note="Taka-amylase A (Taa-GI) precursor"
            /number=1
            intron
            773..827
            exon
            828..868
            /note="Taka-amylase A precursor"
            /number=2
            intron
            869..954
```

```
/note="Taa-G1 intron B"
955..1070
/note="Taka-amylase A precursor"
/number=3
1071..1139
/note="Taa-G1 intron C"
1140..1248
/note="Taka-amylase A precursor"
/number=4
1249..1316
/note="Taa-G1 intron D"
1317..1545
/note="Taka-amylase A precursor"
/number=5
1546..1602
/note="Taa-G1 intron E"
1603..1765
/note="Taka-amylase A precursor"
/number=6
1766..1829
/note="Taa-G1 intron F"
1830..1976
/note="Taka-amylase A precursor"
/number=7
1977..2040
/note="Taa-G1 intron G"
2041..2281
/note="Taka-amylase A precursor"
/number=8
2282..2359
/note="Taa-G1 intron H"
2360..2647
/note="Taka-amylase A precursor"
/number=9
2784..2789
polya_signal
ORIGIN
Query Match 57.0%; Score 989; DB 8; Length 2935;
Best Local Similarity 74.7%; Pred. No. 2.6e-228; Indels 541; Gaps 8;
Matches 1613; Conservative 0; Mismatches 5;
QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTGG 60
DB 557 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTGG 616
QY 61 CGTGGTGTCTCTATTTCTGTAGGCGCTTCAGGTGCGGCACTGCTTTGGCTGCAAGC 120
DB 617 CGTGGTGTCTCTATTTCTGTAGGCGCTTCAGGTGCGGCACTGCTTTGGCTGCAAGC 676
QY 121 CTGGGACTGGCGATCGCAATCCATTTATTTCTCTCAGGATCGATTTGCAAGGACGG 180
DB 677 CTGGGACTGGCGATCGCAATCCATTTATTTCTCTCAGGATCGATTTGCAAGGACGG 736
QY 181 ATGGGTTCGACGACTGGCACTTGTAACTGCTGATC----- 216
DB 737 ATGGGTTCGACGACTGGCACTTGTAACTGCGGATCGGGTGTGTTACCTACTAGCTT 796
QY 217 -----AGAAATACCTGTGGTGGAAACATGGCGAGGC 245
DB 797 TCAGAAAGAGGAATGTAAACTGACTTGTATAGAAATCTGTGGTGGAAACATGGCGAGGC 856
QY 246 ATCATCGACA----- 255
DB 857 ATCATCGACAAGGTAAATTGCCCTTTATCAAAAAAAGAGGAAAGCAAGAA 916
QY 256 -----AGTTGGACTATATCCAGGGAATGG 279
DB 917 ATAAATAAAGAACTCTAGTCCCTAAACCATCACATAGTTGGACTATATCCAGGGAATGG 976
QY 280 GCTTACAGCCATCTGGATCACCCCGGTTACAGCCAGCTGCCCGAGCAACCGCATATG 339
DB 977 GCTTACAGCCATCTGGATCACCCCGGTTACAGCCAGCTGCCCGAGCAACCGCATATG 1036
```

```
QY 340 GAGATGCCCTACCATGGCTACTGGCAGCAGGATAT----- 373
DB 1037 GAGATGCCCTACCATGGCTACTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTAC 1096
QY 374 -----ATACTCTCTGAACGAAA 390
DB 1097 CTGTCACTCTTTTACATCAATATGAACTAACTTGATGGTTTATAGATACTCTCTGAACGAAA 1156
QY 391 ACTACGGCACTGCAGATGACTTTGAAGGCGCTCTCTTTCCGCCCTTTCATGAGAGGGGATGTT 450
DB 1157 ACTACGGCACTGCAGATGACTTTGAAGGCGCTCTCTTTCCGCCCTTTCATGAGAGGGGATGTT 1216
QY 451 ATCTTATGTCGATGTCGTTGCTAAACCATAT----- 481
DB 1217 ATCTTATGTCGATGTCGTTGCTAAACCATATGTCGTCGTGTCCTTTGCAACTGACTTCGC 1276
QY 482 -----GGCTATGATGGAGCGGCTAG 502
DB 1277 GGATATGGTTTCATTTTCAGTACTGACAATGAGTAATATCAGGGCTATGATGGAGCGGCTAG 1336
QY 503 CTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCCAAGACTACTTCCACCCGTTCTG 562
DB 1337 CTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCCAAGACTACTTCCACCCGTTCTG 1396
QY 563 TTTCAATCAAACTATGAAGATCAGACTCAGGTTGAGGATGCTGGCTAGGAGATAACAC 622
DB 1397 TCTCAATCAAACTATGAAGATCAGACTCAGGTTGAGGATGCTGGCTAGGAGATAACAC 1456
QY 623 TGTCTCTTGGCTGATCTCGATACCAAGGATGTCGTAAGATGATGATGATGATGATGATG 682
DB 1457 TGTCTCTTGGCTGATCTCGATACCAAGGATGTCGTAAGATGATGATGATGATGATGATG 1516
QY 683 GGTGGGATCATTTGGTATCGAACTACTCCA----- 711
DB 1517 GGTGGGATCATTTGGTATCGAACTACTCGAGTAAGATATTTCTCCCTCATTTCTACAACTGG 1576
QY 712 -----TTGACGGCCTCGGTATCGACACAGTAAACACGT 745
DB 1577 GCTGATCGATGATCTTACGAAATCAGTTGACGGCCTCGGTATCGACACAGTAAACACGT 1636
QY 746 CCAGAAGGACTTCTGGCCCGGGGTACAAAGCCGCGGCGTGTACTGTATGCGCGAGGT 805
DB 1637 CCAGAAGGACTTCTGGCCCGGGGTACAAAGCCGCGGCGTGTACTGTATGCGCGAGGT 1696
QY 806 GCTCCAGCGTGTATCGGCGCTACACTTGTCTCCACAGAACGTCATGCGAGCGGTACTGAA 865
DB 1697 GCTCCAGCGTGTATCGGCGCTACACTTGTCTCCACAGAACGTCATGCGAGCGGTACTGAA 1756
QY 866 CTATCCCAT----- 874
DB 1757 CTATCCCATGTATGGTTCTCCCAACCATGAGCCTTCTTCAAGTCTCATCTCCTAACGAA 1816
QY 875 -----TTATATCCACTCTCTCAACGCTTCAAGTCAACCTCGGCGAGCATGG 921
DB 1817 AGCGTAAACCACTGTTACTATCCACTCTCAACGCTTCAAGTCAACCTCGGCGAGCATGC 1876
QY 922 AGGACTCTTACAACTATGATCAACCGTCGATCAATCCGACTGTCCGACTCAACACTCCTGG 981
DB 1877 AGGACTCTTACAACTATGATCAACCGTCGATCAATCCGACTGTCCGACTCAACACTCCTGG 1936
QY 982 GCACATTCGTCGAGAACCCACGACCAACCCACGCTTGGCTTC----- 1021
DB 1937 GCACATTCGTCGAGAACCCACGACCAACCCACGCTTGGCTTCGTTAAAGTCTTCCCTTTATTT 1996
QY 1022 -----TTACACCAACGACATA 1037
DB 1997 TCGTTTCCCAATTTCCACACAGAACCCCACTAACAGAGCAAAAGTTTACACCAACGACATA 2056
QY 1038 GCCCTCGCCCAAGACGTGCGAGCATTCATCTCCCAAGCGGAAATCCCATCATCTCTAC 1097
DB 2057 GCCCTCGCCCAAGACGTGCGAGCATTCATCTCCCAAGCGGAAATCCCATCATCTCTAC 2116
QY 1098 GCCGGCCAAGAACAGCACTTACCGCGCGGAAACGACCCCGGAAACCGGAAAGCAACCTGG 1157
```



```
Db 2117 GC CGG CCAAG AAG CAG CACTAG CGCGG GGA AAG CAG CCG CCG AAG CCG AAG CACTGG 2176
QY 1158 CTCTCGGGCTACCGACCGACGAGCGCTCTACAAGTTAATTCCTCCGCGAAGCGCAATC 1217
Db 2177 GCTTCGGGCTACCGACCGACGAGCGCTCTACAAGTTAATTCCTCCGCGAAGCGCAATC 2236
QY 1218 CGAAGCTATGCCATAGCAAGATACAGGATTCGTGACCTTACA----- 1260
Db 2237 CGAAGCTATGCCATAGCAAGATACAGGATTCGTGACCTTACAAGGTAAGCAACCTCT 2296
QY 1261 ----- 1260
Db 2297 AAGCATACCTTAATGGCTATCTCTCAGAGTATCTGACACAGACTAATCACTGGCAATA 2356
QY 1261 -AGAACTGGCCCTCTACAAGACGACACACAGATCGCCATCGCGAAGGCGACAGATGGG 1319
Db 2357 CAGAACTGGCCCTCTACAAGACGACACACAGATCGCCATCGCGAAGGCGACAGATGGG 2416
QY 1320 TCGCAGATCGTGACTATCTTGTCCACAGAGGTGCTTCGGGTGATTCGTATACCTCTCC 1379
Db 2417 TCGCAGATCGTGACTATCTTGTCCACAGAGGTGCTTCGGGTGATTCGTATACCTCTCC 2476
QY 1380 TTGAGTGGTGGGTTACAGCGCGCGACCAATTGACGAGGTCTATTGGCTGCAGACC 1439
Db 2477 TTGAGTGGTGGGTTACAGCGCGCGACCAATTGACGAGGTCTATTGGCTGCAGACC 2536
QY 1440 GTGACGGTGGTTCGGATGGAATGTCCCTGTTCTATGGCAGGTGGGCTACCTAGGGTA 1499
Db 2537 GTGACGGTGGTTCGGATGGAATGTCCCTGTTCTATGGCAGGTGGGCTACCTAGGGTA 2596
QY 1500 TTGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGTCGTGAAGGGTGGAG 1559
Db 2597 TTGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGTCGTGAAGGGTGGAG 2656
QY 1560 AGTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGGTTGATGTA 1618
Db 2657 AGTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGGTTGATGTA 2715

RESULT 12
AF139925
LOCUS Aspergillus flavus alpha-amylase (amy1) gene, complete cds.
DEFINITION Aspergillus flavus alpha-amylase (amy1) gene, complete cds.
ACCESSION AF139925
VERSION AF139925.1 GI:6492298
KEYWORDS
SOURCE Aspergillus flavus
ORGANISM Aspergillus flavus
REFERENCE 1 (bases 1 to 2970)
AUTHORS Fakhoury,A.M. and Woloshuk,C.P.
TITLE Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
aflatoxin biosynthesis in maize kernels
JOURNAL Phytopathology 89, 908-914 (1999)
AUTHORS Fakhoury,A.M. and Woloshuk,C.P.
TITLE Direct Submission
JOURNAL Submitted (01-APR-1999) Botany and Plant Pathology, Purdue
University, 1155 Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
source
1. .2970
/organism="Aspergillus flavus"
/mol_type="genomic DNA"
/strain="86-10D"
/db_xref="taxon:5059"
<433. .>2481
/gene="amy1"
join(<433. .600,656. .694,785. .900,970. .1078,1147. .1375,
1434. .1596,1662. .1808,1874. .2114,2194. .2481)
/gene="amy1"
/product="alpha-amylase"
```

```
CDS
join(433. .600,656. .694,785. .900,970. .1078,1147. .1375,
1434. .1596,1662. .1808,1874. .2114,2194. .2481)
/gene="amy1"
/EC_number="3.2.1.1"
/codon_start=1
/product="alpha-amylase"
/protein_id="AAP14264.1"
/db_xref="GI:6492298"
/translation="MMVAWSLSFLYGLQVAAPALAAATPADRWSQSIFYLLTDRPARTD
GSTTATCNADRYKCGTGWQIGIDKLDYIQGMGFTAIWITPVTYQAIPQTTAYGDYVHG
YWOQDIYSLNLEKADLKSALSHRGMTYLVVDVANHMGYDCAGSSVDYVFKP
FSSDYFHEPCLIQNYEDOTVEDCWLGNTVSLDPLDTTKDVVKNEWDMVGLSVSN
YSIDGLRIDTVKHVKQDFWPKYKAGVYICIEVLGDGPAYTCPYQNMVGLVNYPIY
YELNARFSTGSMDDLNMINTVRSDCPSILGTFVENHONPRFASITNDIALAKN
YAFILINDGIPILYAGQHQYAGNDPANREATWLSGYPTTSELKYLJASANAIRNY
AISKDTGFYTKNWIYIKDDTTIAMRKGTGDSQIVTILSNKAGSGSYTSLSGAGYT
AQQLTEVIGCTTVTVGSDGNVPMAGGLPRVLVPTKLAGSKICSSS"

ORIGIN
Query Match 56.9%; Score 985.8; DB 8; Length 2970;
Best Local Similarity 74.6%; Pred. No. 1.6e-227;
Matches 1616; Conservative 0; Mismatches 2; Indels 549; Gaps 8;

QY 1 TCACATCAAGCTCTCCCTTCTCTGACGAGTAAACCCACAGAGGCGATTTATGATGTCG 60
Db 383 TCACATCAAGCTCTCCCTTCTCTGACGAGTAAACCCACAGAGGCGATTTATGATGTCG 442
QY 61 CGTGGTGGTCTCTATTCTGACGGCTTTCAGTCCGCGCACCTGCTTGGCTGCAACGC 120
Db 443 CGTGGTGGTCTCTATTCTGACGGCTTTCAGTCCGCGCACCTGCTTGGCTGCAACGC 502
QY 121 CTGCGGACTGGCGATCGCAATCAATTTATTTCTCTCAGGATCGATTTGCAAGGACGG 180
Db 503 CTGCGGACTGGCGATCGCAATCAATTTATTTCTCTCAGGATCGATTTGCAAGGACGG 562
QY 181 ATGGGTGCGAGCTCGGACTTGTATCTGCGGATC----- 216
Db 563 ATGGGTGCGAGCTCGGACTTGTATCTGCGGATCGGATGTTGTTGCTTACTAGCTT 622
QY 217 -----AGAAATACTGTGGTGGAAACATCGCAGGGC 245
Db 623 TCAGAAAGAGGAATGTAAACTGACTGTATAGTAATCTGTGTGGACATCGCAGGGC 682
QY 246 ATCATCGACA----- 255
Db 683 ATCATCGACAAGTAAATTTGCCCTTTATCAAAAAAAAAAAAAAGAAAGCAGAG 742
QY 256 -----AGTTGACATATATCCAGGGA 275
Db 743 AAAAAATAAATAAAGAACTCTAGTCTTAACCATCATATAGTGTGGACTATATCCAGGGA 802
QY 276 ATGGGCTTCACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCGACACCGCA 335
Db 803 ATGGGCTTCACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCGACACCGCA 862
QY 336 TATGAGATGCTCAACATCGGCTACTGGCAGCAGGATAT----- 373
Db 863 TATGAGATGCTCAACATCGGCTACTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATAT 922
QY 374 -----ATACTCTCTGAAC 386
Db 923 CTACCTGTGATCTTTTACATCAATATGAACCTTGTATGTTTATAGTACTCTCTGAAC 982
QY 387 GAAACTACCGCAGCTGCGAGATGACTTGAAGGGCTCTCTTCGGGCCCTTCATGAGAGGGG 446
Db 983 GAAACTACCGCAGCTGCGAGATGACTTGAAGGGCTCTCTTCGGGCCCTTCATGAGAGGGG 1042
QY 447 ATGTATCTTATGTCGATGTTGGTTGCTAACCATAT----- 481
Db 1043 ATGTATCTTATGTCGATGTTGGTTGCTAACCATATGTTTCGTTGCTTGTGCAACTGACT 1102
QY 482 -----GGCTATGATGGAGCGG 498
|||||
```

Db 1103 TCGCGGATATGTTTCATTTTCAGTACTGACAAATGAGTAATATCAGGGCTATGATGAGCGG 1162
QY 499 GTAGCTCAGTCAGTACAGTGTGTTTAAACCGTTCAGTTCCTCCCAAGACTACTTCCACCCGT 558
Db 1163 GTAGCTCAGTCAGTGTGTTTAAACCGTTCAGTTCCTCCCAAGACTACTTCCACCCGT 1222
QY 559 TCTGTTTCATTCAAAACCTATGAAGATCAGACTCAGGTTTGAGGATTCGCTGGCTAGGAGATA 618
Db 1223 TCTGTTTCATTCAAAACCTATGAAGATCAGACTCAGGTTTGAGGATTCGCTGGCTAGGAGATA 1282
QY 619 ACAGTGTCTCTGCTGCTGATCTCGATACCAACCAAGATGTGTCAAGATGAATGATGATGAG 678
Db 1283 ACAGTGTCTCTGCTGCTGATCTCGATACCAACCAAGATGTGTCAAGATGAATGATGATGAG 1342
QY 679 ACTGGGTGGGATCATTTGGTATCGAACTACTCCA----- 711
Db 1343 ACTGGGTGGGATCATTTGGTATCGAACTACTCCAAGATGATTTCCCTCCCTCATTTCTACAA 1402
QY 712 -----TTGACGGCTCCCGTATCGACACAGTAAAA 740
Db 1403 CTTGGCTGATCGATGATCTTACGAAATCAGTTGACGGCTCCGTTATCGACACAGTAAAA 1462
QY 741 CACGTCACAGAGACTTCTGGCCCGGTACAAACCAAGCGGAGCGGTGATGTTATCGGC 800
Db 1463 CACGTCACAGAGACTTCTGGCCCGGTACAAACCAAGCGGAGCGGTGATGTTATCGGC 1522
QY 801 GAGGTGCTCGACGGTGATCCGGCTACACTCTGCTCTACCAAGCTCATGATGACGCGCTA 860
Db 1523 GAGGTGCTCGACGGTGATCCGGCTACACTCTGCTCTACCAAGCTCATGATGACGCGCTA 1582
QY 861 CTGAATATTCCTAT----- 874
Db 1583 CTGAATATTCCTATGTTCTCCACCATGAGCCTTCTTGGCAAGCTCATCTCCTA 1642
QY 875 -----TTACTATCGACTCTCTCAACCGCTTCAAGTCAACCTCGGCA 915
Db 1643 ACGAAACGGCTAAAAACAGTATCTATCTCACTCTCAACCGCTTCAAGTCAACCTCGGCA 1702
QY 916 GCATGACGACCTCTACACATGATCAACACCGTCAAAATCCGACTGTCAGACTCAACAC 975
Db 1703 GCATGACGACCTCTACACATGATCAACACCGTCAAAATCCGACTGTCAGACTCAACAC 1762
QY 976 TCTGGGACATTCGTCGAGAACCCAGCAACCCAGCGTTTCGCTTC----- 1021
Db 1763 TCTGGGACATTCGTCGAGAACCCAGCAACCCAGCGTTTCGCTTCGTAAGTCTTCCCTT 1822
QY 1022 -----TTACACCAA 1030
Db 1823 TTATTTTTCGGTTTCCCAATTTCCACAGAACCCCATCTAACAGAGCAAAAGTTACACCAA 1882
QY 1031 CGACATAGCCCTCGCCCAAGACGTGCGACATTCATCTCATCTCAACGCGGAATCCGCAT 1090
Db 1883 CGACATAGCCCTCGCCCAAGACGTGCGACATTCATCTCATCTCAACGCGGAATCCGCAT 1942
QY 1091 CATCTACGCGCGCCAGAACAGCACTATACCGCGCGGAAACGACCCCGCGAAACCGGAAAGC 1150
Db 1943 CATCTACGCGCGCCAGAACAGCACTATACCGCGCGGAAACGACCCCGCGAAACCGGAAAGC 2002
QY 1151 AACCTGGCTCTCGGGTACCCGACGACAGCGAGCTGTACAAGTTAATTTGCTTCGCGGAA 1210
Db 2003 AACCTGGCTCTCGGGTACCCGACGACAGCGAGCTGTACAAGTTAATTTGCTTCGCGGAA 2062
QY 1211 CGCAATCCGGAATATGCTATGCAAGATACAGGATTCGTAAGCTTCA----- 1260
Db 2063 CGCAATCCGGAATATGCTATGCAAGATACAGGATTCGTAAGCTTCAAGGTAAGCAC 2122
QY 1261 ----- 1260
Db 2123 AACCTTAAGCATACCTAATGCTTCTTTCAGAGTATCTGACACAGAGACTAATCAC 2182
QY 1261 -----AGAACTGGCCCATCTACAAAGACGACACAAAGATCGCCATCGCGCAAGGCA 1311
Db 2183 TGGCAATACAGAACTGGCCCATCTACAAAGACGACACAAAGATCGCCATCGCGCAAGGCA 2242

QY 1312 CAGATGGTTCGACAGATCGTACTATCTTTGTCCAAACAGGGTCTCTTCGGGTGATTCGTATA 1371
Db 2243 CAGATGGTTCGACAGATCGTACTATCTTTGTCCAAACAGGGTCTCTTCGGGTGATTCGTATA 2302
QY 1372 CCTCTCCTTCGAGTGGTTCGGGTTACACAGCGGGCAGCAATTTGACGAGGTCAATTGGCT 1431
Db 2303 CCTCTCCTTCGAGTGGTTCGGGTTACACAGCGGGCAGCAATTTGACGAGGTCAATTGGCT 2362
QY 1432 GCACGACGTCGACGGTTCGGTTCGGATGGAATGTCCCTGTTCTCTATGGCAGGTGGCTAC 1491
Db 2363 GCACGACGTCGACGGTTCGGTTCGGATGGAATGTCCCTGTTCTCTATGGCAGGTGGCTAC 2422
QY 1492 CTAGGTTATTTGATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTACTCGTGAA 1551
Db 2423 CTAGGTTATTTGATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTACTCGTGAA 2482
QY 1552 GGTGGAGAGTATATGATGGTACTGCTATTCATCTGGCATTTGGACAGTGTGAGTT 1611
Db 2483 GGTGGAGAGTATATGATGGTACTGCTATTCATCTGGCATTTGGACAGTGTGAGTT 2542
QY 1612 TGATGTA 1618
Db 2543 TGATGTA 2549

RESULT 13
ANAMYB ANAMYB 2443 bp DNA linear PLN 01-DEC-1992
LOCUS A.niger amyB gene for alpha-amylase.
DEFINITION X52756
ACCESSION X52756
VERSION 1
KEYWORDS alpha-amylase.
SOURCE Aspergillus niger
ORGANISM Aspergillus niger
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 2443)
AUTHORS Korman,D.R., Bayliss,F.T., Barnett,C.C., Carmona,C.L., Kodama,K.H.,
Royer,T.J., Thompson,S.A., Ward,M., Wilson,L.J. and Berka,R.M.
TITLE Cloning, characterization, and expression of two alpha-amylase
genes from Aspergillus niger var. awamori
JOURNAL Curr. Genet. 17 (3), 203-212 (1990)
MEDLINE 90254827
PUBMED 2340591
FEATURES
Location/Qualifiers
1..2443
/organism="Aspergillus niger"
/mol_type="genomic DNA"
/strain="var.awamori, UVK143f"
/db_xref="taxon:5061"
11..13
101..106
201..2244
/gene="amyB"
join(201..368,424..462,548..663,733..841,910..1138,
1197..1359,1425..1571,1637..1877,1957..2244)
/gene="amyB"
/EC_number="3.2.1.1"
/function="hydrolysis of alpha-1,4-glucosidic bonds in
polyglucosans"
/note="1, 4-alpha-D-glucan 4-glucanohydrolase"
/codon_start=1
/product="alpha-amylase-precursor"
/protein_id="CAA36967.1"
/db_xref="GI:2326"
/db_xref="GOA:Q02906"
/translation="MMVAWWSLFLYLQVAPALAAATPADMRSQSTYFLLLDRFARTD
GSTATCATADQKYCGTWTQGIIDKLDTIQMGFTAIWITPVTAOLPOTTAAGDAVHG
YQQDIYSINENGTGADDLKALSSALHGERMYLWVDVNVANHNGYDAGSSVDYSVFKP
FSSQDYFHPFCFIQNYEDQVDCWLGNTVSLPLDITKDVVKNEWYDWVGLSVSN
YSIDLRTITVKHVKQDFWPGYNKAAGVYCTGEVLDDGDPAYTCTPQNVNMDVGLNPIY

CAAT_signal
TATA_signal
gene

CDS

YPLLAFKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHNDPRASVYNDIALAKN
VAAPILINDSIPPIYAGOEYHAGNDPANREBATWLSGYPTDSELKLIASRAIRNY
AISKDGFVYIKWPIYKODITIPMRKGTGDSQIVTILSNKGASGDSYTLUSLGAGYT
AGQOLTEVIGCTTIVTGVSDGNVPPVMAGGLPRVLYPTEKLAGSKICSSS"

sig_peptide

/gene="amyB"

/notes="alpha-amy-lase"

join(264..358,424..462,548..663,733..841,910..1138,

1197..1359,1425..1571,1637..1877,1957..2241)

/gene="amyB"

/product="alpha-amy-lase"

/EC number="3.2.1.1"

/functions="hydrolysis of alpha-1,4-glucosidic bonds in

polyglucosans"

/notes="1,4-alpha-D-glucan 4-glucanohydrolase"

201..368

/gene="amyB"

/number=1

369..423

/gene="amyB"

/number=1

424..462

/gene="amyB"

/number=2

463..547

/gene="amyB"

/number=2

548..663

/gene="amyB"

/number=3

664..732

/gene="amyB"

/number=3

733..841

/gene="amyB"

/number=4

842..909

/gene="amyB"

/number=4

910..1138

/gene="amyB"

/number=5

1139..1196

/gene="amyB"

/number=5

1197..1359

/gene="amyB"

/number=6

1360..1424

/gene="amyB"

/number=6

1425..1571

/gene="amyB"

/number=7

1572..1636

/gene="amyB"

/number=7

1637..1877

/gene="amyB"

/number=8

1878..1956

/gene="amyB"

/number=8

1957..2244

/gene="amyB"

/number=9

2336..2341

polyA_site
ORIGIN

Query Match 56.5%; Score 979; DB 8; Length 2443;

Best Local Similarity 74.4%; Pred. No. 6.7e-226;

Matches 1609; Conservative 0; Mismatches 10; Indels 544; Gaps 8;

QY	1	TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCAATTTATGATGGTCG	60
DB	151	TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCAATTTATGATGGTCG	210
QY	61	CGTGGTGGTCTCTATTCTGTACGGCTTTCAGGTCCGGCACCTGTTGGCTGCAACGC	120
DB	211	CGTGGTGGTCTCTATTCTGTACGGCTTTCAGGTCCGGCACCTGTTGGCTGCAACGC	270
QY	121	CTGCGGACTGGCGATCGCAATCAATTTATTTCTCTCAGGATCGATTGTCGAAGACGG	180
DB	271	CTGCGGACTGGCGATCGCAATCAATTTATTTCTCTCAGGATCGATTGTCGAAGACGG	330
QY	181	ATGGGTGCGAGCTGCGACTTGTAACTACGCGATC-----	216
DB	331	ATGGGTGCGAGCTGCGACTTGTAACTACGCGATCAGGTGTGTTGTTACCTACTAGCTT	390
QY	217	-----AGAAATACCTGTGTGGTGGAAACATGCGAGGGC	245
DB	391	TCGAAAAGAGGAATGTAAACTGACTTGATATAGAAATACTGTGTGGTGGAAACATGCGAGGGC	450
QY	246	ATCATCGACA-----	255
DB	451	ATCATCGACAAGTAAATTTGCCCTTTATCAAAAAAAGAGAAAGAGAGAGAAAA	510
QY	256	-----AGTTGGACTATATCCAGGGGAATGGG	280
DB	511	TAAATAAAGAACTCTAGTCTTAACCATCATAGTTGGACTATATCCAGGGGAATGGG	570
QY	281	CTTCACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCCAGACACCGCATATGG	340
DB	571	CTTCACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCCAGACACCGCATATGG	630
QY	341	AGATGCTTACCATGGCTACTGGCAGCAGGATAT-----	373
DB	631	AGATGCTTACCATGGCTACTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC	690
QY	374	-----ATACTCTCTGAACGAAAA	391
DB	691	TGTCATCTTTTACATCAATATGAACCTTAACCTTGATGGTTTTAGATACTCTCTGAACGAAAA	750
QY	392	CTACGGCACTGAGATGACTTGAAGCGCTCTCTCGGCCCTTCATGAGAGGGGATGTA	451
DB	751	CTACGGCACTGAGATGACTTGAAGCGCTCTCTCGGCCCTTCATGAGAGGGGATGTA	810
QY	452	TCTTATGGTCGATGGTGTCTAAACCATAT-----	481
DB	811	TCTTATGGTCGATGGTGTCTAAACCATATGGTTCGTGGTCTCTTTCGAACCTGCTCGCG	870
QY	482	-----GGGCTATGATGAGCGGGTAGC	503
DB	871	GATATGGTTCATTTTCAGTACTGACAATGAGTAATATACAGGGCTATGATGGAGCGGGTAGC	930
QY	504	TCAGTCGATTACAGTGTGTTTAAACCGTTCAGTTCCTCCAGAGACTTCCACCGGTTCTGT	563
DB	931	TCAGTCGATTACAGTGTGTTTAAACCGTTCAGTTCCTCCAGAGACTTCCACCGGTTCTGT	990
QY	564	TTTCAATTTCAAACTATGAAAGATCAGACTCAGGTGTGAGGATTTGCTGGCTAGAGATAACACT	623
DB	991	TTTCAATTTCAAACTATGAAAGATCAGACTCAGGTGTGAGGATTTGCTGGCTAGAGATAACACT	1050
QY	624	GTCTCTTTCCTGATCTCGATACCAAGGATGTGGTCAAGAATGAATGGTACGACTGG	683
DB	1051	GTCTCTTTCCTGATCTCGATACCAAGGATGTGGTCAAGAATGAATGGTACGACTGG	1110
QY	684	GTGGGATCATTTGGTATCGAACTACTCCA-----	711
DB	1111	GTGGGATCATTTGGTATCGAACTACTCCAAGTAAGATAATTTCTCCCTCAATTTCTACAACTGG	1170
QY	712	-----TTGACCGCTCTCGGTATCGACACAGTAAAAACAGCT	745
DB	1171	CTGATCGATGATCTTACGAAATCAGTTGACGGCTCTCGGTATCGACACAGTAAAAACAGCT	1230
QY	746	CCAGAAGGACTTCTGGCCCGGGGTACAAAGCCGCGGTGTACTGTATCGGCGAGGT	805

Db	1231	CGAAGGACTTCTGGCCCGGTACAAACAAGCGCAGCGGTACTGTATCGCGAGGT	1290
QY	806	GCTCGACGGTGATCCGGCCCTACACTTGTCCTTACCAGACGTCATGGACGGGTACTGAA	865
Db	1291	GCTCGACGGTGATCCGGCCCTACACTTGTCCTTACCAGAACGTCATGGACGGGTACTGAA	1350
QY	866	CTATCCCAT-----	874
Db	1351	CTATCCCATGTATGGTTCTCCCAACATGAGCCTTCTTCCAAGTCTCATCTCTCTAACGAA	1410
QY	875	-----TTACTATCACTCTCTCAACGCTTCAAGTCAACTCCGGCAGCATG	920
Db	1411	ACGGCTAAACAGTTACTATCACTCTCTCAACGCTTCAAGTCAACTCCGGCAGCATG	1470
QY	921	GACGACCTTCAACATGATCAACACCGTCAAAATCCGACTGCCAGCTCAACATCCCTG	980
Db	1471	GACGACCTTCAACATGATCAACACCGTCAAAATCCGACTGCCAGCTCAACATCCCTG	1530
QY	981	GGCACATTCGTGGAGAACCAACGACCAACCCACGTTGGCTTC-----	1021
Db	1531	GGCACATTCGTGGAGAACCAACGACCAACCCACGTTGGCTTCGTAAGTCTTCCCTTTTATT	1590
QY	1022	-----TTACACCAACGACA	1035
Db	1591	TTCCGTTCCCAATTTCCACACAGAACCCCACTTAACAGAGCAAAAGTTACACCAACGACA	1650
QY	1036	TAGCCCTCGCAAGACGTCGAGGATTCATCTCTCAACGAGGAATCCCCATCATCT	1095
Db	1651	TAGCCCTCGCAAGACGTCGAGGATTCATCTCTCAACGAGGAATCCCCATCATCT	1710
QY	1096	ACGCCGGCCAAAGACAGCACTACGCGCGGGAACGACCCCGCAACCGCAAGCAACT	1155
Db	1711	ACGCCGGCCAAAGACAGCACTACGCGCGGGAACGACCCCGCAACCGCAAGCAACT	1770
QY	1156	GGCTCTCGGGCTACCCGACCGACAGCGACTGTACAAAGTTAATGCTCCGCGAAACGCAA	1215
Db	1771	GGCTCTCGGGCTACCCGACCGACAGCGACTGTACAAAGTTAATGCTCCGCGAAACGCAA	1830
QY	1216	TCGGNAACTATGCCATTAGCAAGATACAGGATTCGTGACCTACA-----	1260
Db	1831	TCGGNAACTATGCCATTAGCAAGATACAGGATTCGTGACCTACAAGGTAAGCACAACCT	1890
QY	1261	-----	1260
Db	1891	CTAAGCATACCTAATGGCTATCTCAGATATCTGACACAGAGACTAATCATCTGGCA	1950
QY	1261	----AGAACTGGCCCATCTACAAAGACGACACAAACGATCGCATCGCAAGGGCAGAT	1316
Db	1951	ATACAGAACTGGCCCATCTACAAAGACGACACAAACGATCCCGATCGCAAGGGCAGAT	2010
QY	1317	GGTCCGACATCGTGACTATCTGTCCACAAAGGGTGTTCGGGTGATTCGTATACCCCTC	1376
Db	2011	GGTCCGACATCGTGACTATCTGTCCACAAAGGGTGTTCGGGTGATTCGTATACCCCTC	2070
QY	1377	TCCTTGAGTGGTGGGGTTTACACAGCGCGCCAGCAATTCACGAGGTCAATTCGCTGCAG	1436
Db	2071	TCCTTGAGTGGTGGGGTTTACACAGCGCGCCAGCAATTCACGAGGTCAATTCGCTGCAG	2130
QY	1437	ACCGTGACCGGTTGGTTCGGATGAAATGTGCCCTGTTCTTATGGCAGGTGGGCTACCTAGG	1496
Db	2131	ACCGTGACCGGTTGGTTCGGATGAAATGTGCCCTGTTCTTATGGCAGGTGGGCTACCTAGG	2190
QY	1497	GTATTGTATCCGACTGAGAAAGTTGGCAGTAGCAGATCTGTAGTAGCTCGTAAGGGTG	1556
Db	2191	GTATTGTATCCGACTGAGAAAGTTGGCAGTAGCAGATCTGTAGTAGCTCGTAAGGGTG	2250
QY	1557	GAGAGTATATGATGGTACTTATCAATCTGGCATTTGGACAGTGGTGTGAGTTTGTATG	1616
Db	2251	GAGAGTATATGATGGTACTTATCAATCTGGCATTTGGACAGTGGTGTGAGTTTGTATG	2310
QY	1617	TAC 1619	

Db	2311	GAC 2313
RESULT 14		
ANAMYA	ANAMYA	
LOCUS	A.niger amyA gene for alpha-amylase.	2520 bp DNA linear PLN 01-DEC-1992
DEFINITION	X52755	
ACCESSION	X52755	
VERSION	X52755.1 GI:2323	
KEYWORDS	alpha-amylase	
SOURCE	Aspergillus niger	
ORGANISM	Aspergillus niger	
REFERENCE	1 (bases 1 to 2520)	
AUTHORS	Korman,D.R., Bayliss,F.T., Barnett,C.C., Carmona,C.L., Kodama,K.H., Royer,T.J., Thompson,S.A., Ward,M., Wilson,L.J. and Berka,R.M.	
TITLE	Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus niger var. awamori	
JOURNAL	Curr. Genet. 17 (3), 203-212 (1990)	
MEDLINE	90254827	
PUBMED	2340591	
FEATURES	Location/Qualifiers	
source	1..2520	
	/organism="Aspergillus niger"	
	/mol_type="genomic DNA"	
	/strain="var.awamori, UVK143f"	
	/db_xref="taxon:5061"	
CAAT_signal	11..13	
TATA_signal	101..106	
gene	201..2241	
	/gene="amyA"	
CDS	join(201..368,424..462,548..663,733..841,910..1138,1197..1359,1425..1571,1637..1877,1957..2241)	
	/gene="amyA"	
	/EC_number="3.2.1.1"	
	/function="hydrolysis of alpha-1,4-glucosidic bonds in polyglucosans"	
	/note="1, 4-alpha-D-glucan 4-glucanohydrolase"	
	/codon_start=1	
	/product="alpha-amylase-precursor"	
	/protein_id="CAA36966.1"	
	/db_xref="GI:2324"	
	/db_xref="GOA:Q02905"	
	/db_xref="UniProt/Swiss-Prot:Q02905"	
	/translation="MMVAWSLFLXGLQVAPALAAATPADWESQSIYELLTDREPTD GSTATCNTAQKCYCGTQGIIDKLDYIQMGFTAIWITPTAQLPQTATGDAYHG YWQDIYSEINENYGTADDLKALLSALHERGYLMVDVVAANHMGVYDGAGSSVDYISVPKP FSSQDYFHPFCFIQNYEDQTQVEDCWLGDNTVSLDPTTKDVKVXNWDYWGSLVSN YSIDGLRIDTVKHQKDFWPGVKAAGVYCGEVLGDGPAYTCTPTQNVMDGVINTPIY YPLNAPKSTSGSMDLLNMTNVTKSDCPDSTLLGTFFVENHDPNPRFASVTNDIALAKN VAAFTILNDGPIIYAGQEHVAGGNDPANREATWLSGYPDSELYKLIASRNAIRNY AISKDTGYTVYKNWFIYKDDTTIPMRKGTGDSQIVTILSNKASGSDSYTLSISGAGYT AQQUTEVIGCTTITVSGDGNVFPVPMAGGLPRVLYPTEKLAGSKICYG"	
sig_peptide	201..263	
	/gene="amyA"	
	/note="alpha-amylase"	
mat_peptide	join(264..368,424..462,548..663,733..841,910..1138,1197..1359,1425..1571,1637..1877,1957..2238)	
	/gene="amyA"	
	/product="alpha-amylase"	
	/EC_number="3.2.1.1"	
	/function="hydrolysis of alpha-1,4-glucosidic bonds in polyglucosans"	
exon	/note="1,4-alpha-D-glucan 4-glucanohydrolase"	
	201..368	
	/gene="amyA"	
intron	/number=1	
	369..423	
	/gene="amyA"	
exon	/number=1	
	424..462	
	/gene="amyA"	

intron	/number=2 463. .547 /gene="amyA" /number=2 548. .663 /gene="amyA" /number=3 664. .732 /gene="amyA" /number=3 733. .841 /gene="amyA" /number=4 842. .909 /gene="amyA" /number=4 910. .1138 /gene="amyA" /number=5 1139. .1196 /gene="amyA" /number=5 1197. .1359 /gene="amyA" /number=6 1360. .1424 /gene="amyA" /number=6 1425. .1571 /gene="amyA" /number=7 1572. .1636 /gene="amyA" /number=7 1637. .1877 /gene="amyA" /number=8 1878. .1956 /gene="amyA" /number=8 1957. .2241 /gene="amyA" /number=9 2370. .2375	QY	256	-----AGTTGGACTATATCCAGGGAATGGG	280
exon		Db	511	TAAATAAAGAACTCTAGTCTTAACCATCATAGTTGGACTATATCCAGGGAATGGG	570
intron		QY	281	CTTCACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCCAGACACCCCATATGG	340
exon		Db	571	CTTCACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCCAGACACCCCATATGG	630
intron		QY	341	AGATGCTTACCATGGCTACTGGCAGCAGGATAT-	373
exon		Db	631	AGATGCTTACCATGGCTACTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC	690
intron		QY	374	-----ATATCTCTCTGAACGAAAA	391
exon		Db	691	TGTCTATCTTTTACATCAATATGAACCTTAACCTTGATGGTTTTAGATACTCTCTGAACGAAAA	750
intron		QY	392	CTACGGCAGCTGAGATGACTTGAAGCGCTCTCTTCGGGCCCTTCATGAGAGGGGGATGTA	451
exon		Db	751	CTACGGCAGCTGAGATGACTTGAAGCGCTCTCTTCGGGCCCTTCATGAGAGGGGGATGTA	810
intron		QY	452	TCTTATGGTCGATGTGGTTGCTAAACCATAT-	481
exon		Db	811	TCTTATGGTCGATGTGGTTGCTAAACCATATGGTTCGTGGTCTTTTGCACCTGACTTCGCG	870
intron		QY	482	-----GGGCTATGATGGAGCGGGTAGC	503
exon		Db	871	GATATGGTTTCAATTCAGTACTGACAAATAGTAATATCAGGGCTATGATGGAGCGGGTAGC	930
intron		QY	504	TCAGTCGATTACAGTGTGTTTAAACCGTTCAAGTCCCAAGACTACTTCCACCCGTTCTGT	563
exon		Db	931	TCAGTCGATTACAGTGTGTTTAAACCGTTCAAGTCCCAAGACTACTTCCACCCGTTCTGT	990
intron		QY	564	TTCAATTCAAAACTATGAAGATCAGACTCAGGTTGAGGATTTGCTGGCTAGGAGATACACT	623
exon		Db	991	TTCAATTCAAAACTATGAAGATCAGACTCAGGTTGAGGATTTGCTGGCTAGGAGATACACT	1050
intron		QY	624	GTCTCTTCCTGATTCGATACCAAGAGATGGTCAAGAATGAATGGTACGACTGG	683
exon		Db	1051	GTCTCTTCCTGATCTCGATACCAAGAGATGGTCAAGAATGAATGGTACGACTGG	1110
intron		QY	684	GTGGGATCATGGTATCGAACTACTCCAGTAAAGATATTTCTCCCTCATTTCTACAACTGG	711
exon		Db	1111	GTGGGATCATGGTATCGAACTACTCCAGTAAAGATATTTCTCCCTCATTTCTACAACTGG	1170
intron		QY	712	-----TTGACGGCCCTCCGTATCGACACAGTAAACACGT	745
exon		Db	1171	CTGATCGATGATCTTACGAAATCAGTTGACGGCCCTCCGTATCGACACAGTAAACACGT	1230
intron		QY	746	CCAGAAAGGACTTCTGGCCCGGGTACAAACAAAGCCGAGCGGTACTGTATCGCGAGGT	805
exon		Db	1231	CCAGAAAGGACTTCTGGCCCGGGTACAAACAAAGCCGAGCGGTACTGTATCGCGAGGT	1290
intron		QY	806	GCTCGACGCTGATCGGCCCTACACTTGCTCCCTACCAAGACGTCATGGACGGCGTACTGAA	865
exon		Db	1291	GCTCGACGCTGATCGGCCCTACACTTGCTCCCTACCAAGACGTCATGGACGGCGTACTGAA	1350
intron		QY	866	CTATCCCAT-	874
exon		Db	1351	CTATCCCATGATGGTTCCTCCAAACCATGAGCGCTTCTTCAAGTCTCATCTCTTAAGGAA	1410
intron		QY	875	-----TTACTATCCACTCTTCAACGCCCTTCAAGTCAACCTCCGCGAGCATG	920
exon		Db	1411	ACGGCTAAACACGATTACTATCCACTCTTCAAGCCCTTCAAGTCAACCTCCGCGAGCATG	1470
intron		QY	921	GACGACCTCTACACATGATCAACCCGTCAAATCCGACTGTCCAGACTCAACACTCCTG	980
exon		Db	1471	GACGACCTCTACACATGATCAACCCGTCAAATCCGACTGTCCAGACTCAACACTCCTG	1530
intron		QY	981	GGCACAATTCTCGAGAACCCACGACACCCACCGTTTCGCTTC-	1021
exon		Db	1531	GGCACAATTCTCGAGAACCCACGACACCCACCGTTTCGCTTCGTAAGTCTTCCCTTTTATT	1590

Query Match	52.4%;	Score	908.2;	DB	8;	Length	2520;
Best Local Similarity	73.6%;	Pred. No.	9.4e-209;				
Matches	1537;	Conservative	0;	Mismatches	8;	Indels	544;
Gaps							
QY	1	TCACATCAAGCTCTCCCTTCTCTGAACAAATAAAACCCACAGAGGCAATTTATGATGGTCG	60				
Db	151	TCACATCAAGCTCTCCCTTCTCTGAACAAATAAAACCCACAGAGGCAATTTATGATGGTCG	210				
QY	61	CGTGGTGTCTTATTTCTGTAAGGCTTCAGGTGCGGGACCTGCTTTGGCTGCAACGC	120				
Db	211	CGTGGTGTCTTATTTCTGTAAGGCTTCAGGTGCGGGACCTGCTTTGGCTGCAACGC	270				
QY	121	CTGCGGACTGGCGATCGCAATCCATTTATTTCTCTCAGGATCGATTTGCAAGGACGG	180				
Db	271	CTGCGGACTGGCGATCGCAATCCATTTATTTCTCTCAGGATCGATTTGCAAGGACGG	330				
QY	181	ATGGGTGCGAGCTGCGACTTGTAAATACTGCGGATC-	216				
Db	331	ATGGGTGCGAGCTGCGACTTGTAAATACTGCGGATCAGGTGTGTTGTACTACTAGCTT	390				
QY	217	-----AGAAATACGTGGTGGAAACATGGCAGGC	245				
Db	391	TCAGAAAGAGGAATGTAAACTGACTTGTATAGAAATACTGTGGTGGAAACATGGCAGGC	450				
QY	246	ATCATCGACA-	255				
Db	451	ATCATCGACAAGGTAAATTTGCCCTTTTATCAAAAAAAGAGGAAAAAGCAGAGAAAAA	510				

QY	1022	-----TTACACCAACGACA	1035	QY	68	GTCTCTATTTCTGTACGGCCTT	CAGGTCCGGCAGCACTGCTTTGGTGTCAACGCGCTGCGGA	127
Db	1591	TTCCGTTTCCCAATTTCCACACAGAACCCACCTTAA	CAGAGCAAAGTTTACACCAACGACA	1650	Db	3	GACAAATCTTTCTGTTTCTGGCCATTTTCTGGCTACAGCTCTGGCAGCCACGCTGCAGA	62
QY	1036	TAGCCCTCGCCAAAGCTCGCAGCATTTATCATCTCTCAACGACGAATCCCCATCATCT	1095	QY	128	CTGGCGATCGCAATCCATTTATTTCTTCTCAGCGATCGATTTTGCAAGGACGATGGGTC	187	
Db	1651	TAGCCCTCGCCAAAGAACTCGCAGCATTTATCATCTCTCAACGACGAATCCCCATCATCT	1710	Db	63	ATGGCGCTCCCAAGTCGATATATTTCTGCTCACCGATCGCTTTGGCGAACGGAATATTC	122	
QY	1096	ACGCCGGCCAAAGAACAGCACTACGCCGGCGGAAACGACCCCGGAACCCGGAAGCAACCT	1155	QY	188	GACGACTCTGCACTTTGTATATCTGCGGATCAGAAATACTGTGTGTGGAACATGGCAGGGCAT	247	
Db	1711	ACGCCGGCCAAAGAACAGCACTACGCCGGCGGAAACGACCCCGGAACCCGGAAGCAACCT	1770	Db	123	TACCACCTGCTTCTTGTGACTTTGAGCGCTCGGCAATATTCGGGTGGATCTCTGGCAGGGCAT	182	
QY	1156	GGCTCTCGGGCTACCCGACCGACAGCGAGCTGTACAAGTTAAATTTGCTCCGCGAAACGCAA	1215	QY	248	CATCGACAAGTTGGACTATATTCAGGGAATGGGCTTTCACAGCCATCTGGATCACCCCCGT	307	
Db	1771	GGCTCTCGGGCTACCCGACCGACAGCGAGCTGTACAAGTTAAATTTGCTCCGCGAAACGCAA	1830	Db	183	CATCAATCAGCTGGACTATATTTCAAGGAATGGGCTTTTACAGCGATCTGGATCACACCCGT	242	
QY	1216	TCGGAACTATGCGATTAGCAAAAGATACAGGATTTGTCGACCTTACA	1260	QY	308	TACAGCCAGCTGCCCCAGACCCGCGATATGGAGATGCCTACCATGGCTACTTGGGACGA	367	
Db	1831	TCGGAACTATGCGATTAGCAAAAGATACAGGATTTGTCGACCTTACAAGGTAAGCACAACCT	1890	Db	243	AACTGCAAGATCCCCCAAGATACCTGGTTACGGACAGGCATATACGGATACCTTGGCAGCA	302	
QY	1261	-----	1260	QY	368	GGATATATATCTCTTGAAACGAAACTACGGCACTGCAGATGACTTTGAAGGCGCTCTCTTC	427	
Db	1891	CTAAGCATACCTTAATGGGCTATCTTCAGAGTATCTGACACAGAGACTAATCACTGGCA	1950	Db	303	GGAGCGTTATGGCCCTGAACCTCCCATTTATGGTACGCGACAGCATCTCAAAGCTCTGGCTTC	362	
QY	1261	-----AGAACTGGCCCATCTACAAAGACGACACAAACGATCGCCATGCGCAAGGGCACAGAT	1316	QY	428	GGCCCTTCATGAGAGGGGATGTATCTTTATGGTCTGATGTGGTTGTCTTAAACATATGGGCTA	487	
Db	1951	ATACAGAACTGGCCCATCTACAAAGACGACACAAACGATCCCGATGCGCAAGGGCACAGAT	2010	Db	363	AGCTCTTCACTCACGGGGCATGTATCTCATGGTGGAGCTTTGTTGGCCAAATCACATATGGGCCA	422	
QY	1317	GGGTGCGAGATGCTGACTATCTTTGCCAACAGAGGTGCTTGGGTGATTTGTTATACCCCTC	1376	QY	488	TGATGGAGGGGTAGCTCAGTCGATACAGTGTGTTTAAACCGTTTCACTTCCCAAGACTA	547	
Db	2011	GGGTGCGAGATGCTGACTATCTTTGCCAACAGAGGTGCTTGGGTGATTTGTTATACCCCTC	2070	Db	423	CAATGGTATCGGGGAGCTCTGTGGACTACAGTGTTTATAGGCCATTTAAATTCGCAAAAGTA	482	
QY	1377	TCCTTCAGTGTGCGGTTTACACAGCGCGCCAGCAATTCACGAGGTCAATTCGGCTGCACG	1436	QY	548	CTTCCACCGCTTCTGTTTTCATTTCAAAACTATGAAGATCAGACTCAGGTTTGAGGATTGCTG	607	
Db	2071	TCCTTCAGTGTGCGGTTTACACAGCGCGCCAGCAATTCACGAGGTCAATTCGGCTGCACG	2130	Db	483	CTTTTCACAACTCTGTTGGATCTCTGATTACAATAACAGACAAAACGTTGAAGACTGCTG	542	
QY	1437	ACCGTGACGTTGGTTCCGATGAAATGTGCTGTTCTTATGCGCAGGTGGGCTACTAGG	1496	QY	608	GCTAGGAGATAACACTGTCTCTTCGCTGATCTCGATACCAACCAAGGATGTGTCAAGAA	667	
Db	2131	ACCGTGACGTTGGTTCCGATGAAATGTGCTGTTCTTATGCGCAGGTGGGCTACTAGG	2190	Db	543	GCTAGGCGATAAACACCGTTGCTTGGCGATCTTGATACTACCACTACGAGGTTGAAGAA	602	
QY	1497	GTATTGTATCCGACTGAGAAGTTGGCAGTAGCAAGATCTGTAGTAGCT	1545	QY	668	TGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCTTTCAGCGGCTCCGTAT	727	
Db	2191	GTATTGTATCCGACTGAGAAGTTGGCAGTAGCAAGATCTGTATCGGCT	2239	Db	603	TATGTGGTATGACTGGGTGCGAGTCTCTGCTCTTAACTACTTCGCGTCAGCGGCTCCGCGT	662	
RESULT 15				QY	728	CGACACAGTAAACACGTCACAGAGGACTTCTGGCCCGGGTACAAACAAAGCCGAGGCGT	787	
AX711606				Db	663	AGACACAGTCAAGAACGTACAGAGAACTTCTGGCCCGGCTACAACTATGCTTCAGGCGT	722	
LOCUS	AX711606	1485 bp	DNA	linear	PAT 10-APR-2003			
DEFINITION	Sequence 7 from Patent WO03016535.			QY	788	GTACTGTATCGCGAGGTGCTCGACGGTGATTCGGGCTTACACTTTGTCCTTACAGAAACGT	847	
ACCESSION	AX711606			Db	723	GTACTGTATTCGAGAAGTCTTCGATGGGACGCTCATACACCTGTCTTATCAGGAAGA	782	
VERSION	AX711606.1	GI:29787678		QY	848	CATGACCGGCTACTGAACATCTCCATTTACTATCTCCTCTCAACGCTTCAAGTCAAC	907	
KEYWORDS	Aspergillus niger			Db	783	CTTTGACGCGAGTCTTAAATACCCCATGTACTATTCCTCTCTCGGGCTTTTTCGAATCCAC	842	
SOURCE	Aspergillus niger			QY	908	CTCCGGCAGCATGGACGACCTCTACAATGATCAACACCGTCAAAATCCGACTGTCCAGA	967	
ORGANISM	Aspergillus niger			Db	843	CAACGGCAGTATCAGCGACCTCTATAACATGATCAACACCGTGAATATCCACCTTCGAGAGA	902	
REFERENCE	1			QY	968	CTCAACACTCTCTGGGCAATTTGTCGAGAAACACGACAAACCCAGGTTCCG---	1024	
AUTHORS	Maier,D., Stock,A., Wagner,C., Folkers,U., Albermann,K. and Hopper,S.			Db	903	TTCTAGCTTCTAGGGACCTTCGTGCAAAACCCAGATAACCCAGCTTTTGCCAAAGCTA	962	
TITLE	Novel amylases and uses thereof			QY	1025	CHACCAACGACATAGCCCTCGCCAAAGAGCTGCGAGCATTTATCATCTCTCAACGCGAAT	1084	
JOURNAL	Patent: WO 03016535-A 7 27-FEB-2003;			Db	963	CACAAAGCAGCATGTCCCTTAGCCAAAAATTCGCCGCAACATTCATCTATCTGTGTGACGCGCAT	1022	
DSM N.V.	(NL)			QY	1085	CCCCATCATCTACGCCGCGCAAGAACGACACTACGCCGGGGGAAACACACCCCGGAACCG	1144	
FEATURES	Location/Qualifiers			Db	1023	TCCCCATATATACGCCCGGTACAGAACAGCATAATAGCGGGGGTATGACCCCTTACAACCG	1082	
source	1..1485			QY	1145	CGAAGCAACCTCGGTCTCGGGCTACCCGACCGACGAGCGAGCTGTACAAAGTTAATTTGCCTC	1204	
	/organism="Aspergillus niger"			QY				
	/mol_type="unassigned DNA"			Db				
	/db_xref="taxon:5061"			QY				
ORIGIN				QY				
Query Match	43.8%;	Score 759.8;	DB 6;	Length 1485;				
Best Local Similarity	70.0%;	Pred. No. 7.3e-173;						
Matches 1038;	Conservative	0;	Mismatches 442;	Indels 3;	Gaps 1;			

[illegible]

Search completed: February 11, 2005, 21:34:13
Job time : 7426 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2005, 14:14:01 ; Search time 906 Seconds

(without alignments)
11329.836 Million cell updates/sec

Title: US-10-820-200-1

Perfect score: 1734

Sequence: 1 tcatcatcaagctctcccttc.....aaaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn19808:*

2: Geneseqn19908:*

3: Geneseqn20008:*

4: Geneseqn20018:*

5: Geneseqn20018b:*

6: Geneseqn20028:*

7: Geneseqn20028b:*

8: Geneseqn20038:*

9: Geneseqn20038b:*

10: Geneseqn20038c:*

11: Geneseqn20038d:*

12: Geneseqn20048:*

13: Geneseqn20048b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1734	100.0	1734	4	Aaf90208 Nucleotid
2	1707	98.4	1914	3	Aaf12832 Aspergill
3	1063.8	61.3	2337	2	Aaq06786 Taka-amyl
4	1011.2	58.3	3285	8	Acc44572 Alpha-amyl
5	979	56.5	2443	13	Adt89627 Aspergill
6	908.2	52.4	2520	13	Adt89631 Aspergill
7	806	46.5	1389	3	Aaf11252 Aspergill
8	786	45.3	1488	10	Abq80346 A. fumiga
9	759.8	43.8	1485	8	Acc43088 Nucleotid
10	648.2	37.4	1893	10	Abq80348 A. fumiga
11	570	32.9	1956	3	Aaq88712 Aspergill
12	539	25.3	1389	3	Aaf11252 Aspergill
13	378.6	21.8	1404	2	Aaq25079 Alpha-amyl
14	377.4	21.8	4214	1	Aan70916 Sequence
15	375.4	21.6	1404	2	Aaq77665 Variant a
16	375.4	21.6	1404	2	Aaq77668 Variant a
17	374.2	21.6	483	3	Aaf11314 Aspergill
18	373.8	21.6	1404	2	Aaq77666 Variant a
19	373.8	21.6	1404	2	Aaq77667 Variant a
20	369.2	21.3	2289	2	Aaq06388 AMY1 gene

C	21	369.2	21.3	4190	8	ABT14461
C	22	369.2	21.3	4190	8	AA50611
C	23	369.2	21.3	4190	10	ABQ84219
	24	368.8	21.3	605	3	AAF11273
	25	349.6	20.2	2290	1	AAH81525
	26	348	20.1	2290	1	AAH81477
	27	344.2	19.9	1048	3	AAF11253
	28	333.6	19.2	1668	8	ACC43090
	29	317.8	18.3	1644	10	ABQ80350
	30	308	17.8	1926	10	ABQ80345
	31	305.2	17.6	1704	8	ACC43093
	32	303.4	17.5	619	8	ABZ56788
	33	299.8	17.3	3001	8	ACC43082
	34	266.6	15.4	2501	8	ACC43084
	35	266.2	15.4	2250	10	ABQ80347
	36	252.4	14.6	1502	13	ADS60482
	37	251	14.5	1850	8	ACC43089
	38	251	14.5	2511	8	ACC43083
	39	247.4	14.3	480	8	ABZ56641
	40	244.8	14.1	1707	10	ABQ80349
	41	236.4	13.6	619	3	AAF12538
	42	220.6	12.7	292	3	AAF11355
	43	206.6	11.9	1602	3	AAF07664
	44	200.8	11.6	986	8	ABZ51373
	45	196.2	11.3	1539	13	ADS47480

ALIGNMENTS

RESULT 1
AAF90208
ID AAF90208 standard; DNA; 1734 BP.
XX
AC AAF90208;
DT 06-AUG-2001 (first entry)
XX
DE Nucleotide sequence of a fungamyl-like alpha-amylase.
KW Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;
KW alcohol; starch; dough improver; brewing; starch liquification; ss.
XX Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT CDS 54..1550
FT /tag= a
FT /product= "fungamyl-like alpha-amylase"
FT sig_peptide 54..113
FT /tag= b
FT mat_peptide 114..1547
FT /tag= c
XX
XX WO200134784-A1.
PN
XX
PD 17-MAY-2001.
XX
XX
PF 10-NOV-2000; 2000WO-DK000626.
XX
PR 10-NOV-1999; 99DK-00001617.
XX
XX (NOVO) NOVOZYMES AS.
PI Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX
DR WPI; 2001-367478/38.
XX P-PSDB; AAB84206.
PT New variant of Fungamyl-like alpha-amylase, useful for production of
PT maltose syrups, includes mutations that improve stability against heat
PT and acidic pH.
XX

Abt14461 HCV envel
Aad50611 pMP30 vec
Abq84219 Vector pM
Aaf11273 Aspergill
Aan81525 Sequence
Aan81477 5' end of
Aaf11253 Aspergill
Acc43090 Nucleotid
Abq80350 A. fumiga
Abq80345 A. fumiga
Acc43093 Nucleotid
Abz56788 Aspergill
Acc43082 Nucleotid
Acc43084 Nucleotid
Abq80347 A. fumiga
Ads60482 Bacterial
Acc43089 Nucleotid
Acc43083 Nucleotid
Abz56641 Aspergill
Abq80349 A. fumiga
Aaf12538 Aspergill
Aaf11355 Aspergill
Aaf07664 Fusarium
Abz51373 Aspergill
Ads47480 Bacterial

Example 1; Page 40-42; 49pp; English.

The present sequence encodes a fungamyl-like alpha-amylase. The specification describes variants of this fungamyl-like alpha-amylase, which have an alteration in one the amino acid regions 98-110, 150-160, 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion or substitution of an amino acid or an insertion of an amino acid downstream of a particular position. The variants retain alpha-amylase activity, and have better heat stability and/or stability at acidic pH, relative to wild-type enzyme. The variants can therefore be used at higher temperatures (more efficient conversion or faster reaction, and have reduced need for cooling and reduced risk of contamination). The variants may also be used in conjunction with other enzymes, particularly glucoamylase during dextrinisation. The variants are used to produce syrups, particularly of high maltose content, or alcohol, from starch; as dough improver for baked goods; in brewing, to increase fermentability of the wort; and for liquefaction of starch

Sequence 1734 BP; 450 A; 460 C; 425 G; 399 T; 0 U; 0 Other;

Query Match 100.0%; Score 1734; DB 4; Length 1734;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTGG 60
1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTGG 60
61 CDTGTTGGTCTCTATTTCTGTAGGCTTTCAGGTGCGGCACTCTGTTGGCTGCAACGC 120
61 CDTGTTGGTCTCTATTTCTGTAGGCTTTCAGGTGCGGCACTCTGTTGGCTGCAACGC 120
121 CTGCGGACTGGCGATCGCAATCAATTTATTTCTTCTCAAGGATCGATTTGCAAGAGCG 180
121 CTGCGGACTGGCGATCGCAATCAATTTATTTCTTCTCAAGGATCGATTTGCAAGAGCG 180
181 ATGGGTTCGACGACTGCGACTTGTATATCTGCGGATCAGAAATACCTGTTGGTGAACATGGC 240
181 ATGGGTTCGACGACTGCGACTTGTATATCTGCGGATCAGAAATACCTGTTGGTGAACATGGC 240
241 AGGGCATCATCGACAAGTTGGACTATATCCAGGGAATGGGCTTTCAGCCATCTGGATCA 300
241 AGGGCATCATCGACAAGTTGGACTATATCCAGGGAATGGGCTTTCAGCCATCTGGATCA 300
301 CCCCCTTACGCCAGCTGCCCGACACCGCATATGAGATGCGCTACCATGGCTACT 360
301 CCCCCTTACGCCAGCTGCCCGACACCGCATATGAGATGCGCTACCATGGCTACT 360
361 GGCAGCAGGATATATATCTCTGAAACGAAACTACGGCACTGCGATGACTTGAAGGCGC 420
361 GGCAGCAGGATATATATCTCTGAAACGAAACTACGGCACTGCGATGACTTGAAGGCGC 420
421 TCTTTTGGCCCTTATGAGAGGGGATGTATCTTATGTCGATGTTGTTGCTAACCCATA 480
421 TCTTTTGGCCCTTATGAGAGGGGATGTATCTTATGTCGATGTTGTTGCTAACCCATA 480
481 TGGGCTATCATGAGCGGGTAGCTCAGTTCGATTCAGTGTGTTTAAACCGTTTCAGTTCCC 540
481 TGGGCTATCATGAGCGGGTAGCTCAGTTCGATTCAGTGTGTTTAAACCGTTTCAGTTCCC 540
541 AAGACTACTTTCACCCGGTCTGTTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGG 600
541 AAGACTACTTTCACCCGGTCTGTTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGG 600
601 ATTGCTGGCTAGGATTAACACTGTCTCTTGCCTGATTCGATACCAACAGGATGTTGG 660
601 ATTGCTGGCTAGGATTAACACTGTCTCTTGCCTGATTCGATACCAACAGGATGTTGG 660
661 TCAGNATGAATGGTACGACTGGTGGGATCATTTGGTATCGAACTACTCCATTTGACGGCC 720
661 TCAGNATGAATGGTACGACTGGTGGGATCATTTGGTATCGAACTACTCCATTTGACGGCC 720
721 TCCGTATCGACACAGTAAAAACGTCAGAGAGGACTTCTTGGCCGGGTACAAACAGCGG 780

DB TCCGTATCGACACAGTAAAAACAGTCCAGAGGACTTCTTGGCCCGGTACAAACAGCGG 780
QY CAGGCGTCTACTGTATCGGCGAGGTGCTCGAGGTGATCCGSCCTACACTTCTCCCTACC 840
DB CAGGCGTCTACTGTATCGGCGAGGTGCTCGAGGTGATCCGSCCTACACTTCTCCCTACC 840
QY AGAAGCTCATGACGCGGCTACTGAACTATCCCATTTACTTATCCACTCTCTCAACGCTTCA 900
DB AGAAGCTCATGACGCGGCTACTGAACTATCCCATTTACTTATCCACTCTCTCAACGCTTCA 900
QY AGTCAACCTCCGCGAGCATGAGACGCTCTCAACATGATCAACACCGTCAAAATCCGACT 960
DB AGTCAACCTCCGCGAGCATGAGACGCTCTCAACATGATCAACACCGTCAAAATCCGACT 960
QY GTCCAGACTCAACACTCTCTGGGCACTTCTGTCGAGAACCCAGCAACCCACCGTTCGCTT 1020
DB GTCCAGACTCAACACTCTCTGGGCACTTCTGTCGAGAACCCAGCAACCCACCGTTCGCTT 1020
QY CTTTACCAACAGCATAGCCCTCGCAGAGAGCTCGCAGCATTTCAATCTCTCAACGAGC 1080
DB CTTTACCAACAGCATAGCCCTCGCAGAGAGCTCGCAGCATTTCAATCTCTCAACGAGC 1080
QY GAATCCCATCATCTACCGCGGCAAGAACAGCACTACGCGCGGCAAAACGCCCGCA 1140
DB GAATCCCATCATCTACCGCGGCAAGAACAGCACTACGCGCGGCAAAACGCCCGCA 1140
QY ACCGGAAGCACTCTGCTCGGCTACCGGCAAGAACAGCACTACGCGCGGCAAAACGCCCGCA 1200
DB ACCGGAAGCACTCTGCTCGGCTACCGGCAAGAACAGCACTACGCGCGGCAAAACGCCCGCA 1200
QY CTTCCGCAACCAATCCGGAATATGCAATAGCAAGATACAGGATTCGTGACCTACA 1260
DB CTTCCGCAACCAATCCGGAATATGCAATAGCAAGATACAGGATTCGTGACCTACA 1260
QY AGAATCGGCCCATCTACAAAGACGACACAAACGATCGCCATGCGCAAGGSCACAGATGGGT 1320
DB AGAATCGGCCCATCTACAAAGACGACACAAACGATCGCCATGCGCAAGGSCACAGATGGGT 1320
QY CGCAGATCGTACTATCTTGTCAACAGGGTCTTCCGGTATTCGTATACCTCTCTCT 1380
DB CGCAGATCGTACTATCTTGTCAACAGGGTCTTCCGGTATTCGTATACCTCTCTCT 1380
QY TGAGTGTGGGTTCAGATGGAATGTGCTTCTTATGCGAGGTGGCTACCTAGGGTAT 1500
DB TGAGTGTGGGTTCAGATGGAATGTGCTTCTTATGCGAGGTGGCTACCTAGGGTAT 1500
QY TGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTCAAGGGTGGAGA 1560
DB TGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTCAAGGGTGGAGA 1560
QY GTATATGATGGTACTGCTATTCATCTGGCATTTGCAAGTGTGAGTTTGTATGATGATCA 1620
DB GTATATGATGGTACTGCTATTCATCTGGCATTTGCAAGTGTGAGTTTGTATGATGATCA 1620
QY GTTGGAGTGGTACTGCTCTATCCCTTATCTCTCGATTTGTTTTCGAAACCCCTAAATG 1680
DB GTTGGAGTGGTACTGCTCTATCCCTTATCTCTCGATTTGTTTTCGAAACCCCTAAATG 1680
QY CCAAGCAGCTAGTCTATATATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
DB CCAAGCAGCTAGTCTATATATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734

RESULT 2
AAFL12832
ID AAF12832 standard; cDNA; 1914 BP.
XX
AC AAF12832;

XX	13-MAR-2001	(first entry)	121	CTCGGAGTCGGGATCGCAATCCATTTATTTCTCTCAGGATCGATTTGCAAGGACGG	180
DT			123	CTCGGAGTCGGGATCGCAATCCATTTATTTCTCTCAGGATCGATTTGCAAGGACGG	182
DE	Aspergillus oryzae	EST SEQ ID NO:5355.	QY		
XX			QY	ATGGGTCGACGACTGCGACTTGTAACTACGCGATCAGAAATACTGTGTGGAACTATGGC	240
KW	Multiple gene expression; filamentous fungal cell; EST;		DB	ATGGGTCGACGACTGCGACTTGTAACTACGCGATCAGAAATACTGTGTGGAACTATGGC	242
KW	expressed sequence tag; Fusarium venenatum; Aspergillus niger;		QY	AGGGCATCATCGACAAAGTTGGACTATATCCAGGGAATGGCTTTCAGAGCCATCTGGATCA	300
KW	Aspergillus oryzae; Trichoderma reesei; identification; recombination;		DB	AGGGCATCATCGACAAAGTTGGACTATATCCAGGGAATGGCTTTCAGAGCCATCTGGATCA	302
KW	culture condition; environmental stress; spore morphogenesis;		QY	CCCCGTTACAGCCGAGCTGCCAGACCAACGCGATATGGAGATGCTTACATGGGTACT	360
KW	metabolic pathway engineering; catabolic pathway engineering; ss.		DB	CCCCGTTACAGCCGAGCTGCCAGACCAACGCGATATGGAGATGCTTACATGGGTACT	362
XX	Aspergillus oryzae.		QY	GGCAGCAGGATATATATCTCTGAAAGAACTTACGGCACTGCGAGATGACTTGAAGCGC	420
XX	WO200056762-A2.		DB	GGCAGCAGGATATATATCTCTGAAAGAACTTACGGCACTGCGAGATGACTTGAAGCGC	422
XX	28-SEP-2000.		QY	TCTTTTGGCCCTTCATGAGAGGGGATGTATCTTATGTGCGATGTGGTTGCTAAACATA	480
XX	22-MAR-2000; 2000WO-US007781.		DB	TCTTTTGGCCCTTCATGAGAGGGGATGTATCTTATGTGCGATGTGGTTGCTAAACATA	482
XX	22-MAR-1999; 99US-00273623.		QY	TGGGCTATGATGGAGCGGGTAGCTCAGTCGATACAGTGTGTTTAAACCGTTTCAGTTCCC	540
XX	(NOVO) NOVO NORDISK BIOTECH INC.		DB	TGGGCTATGATGGAGCGGGTAGCTCAGTCGATACAGTGTGTTTAAACCGTTTCAGTTCCC	542
XX	(NOVO) NOVO NORDISK AS.		QY	ATAGCTACTTTCACCCGTTCTGTTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG	600
XX	Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;		DB	ATAGCTACTTTCACCCGTTCTGTTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG	602
XX	WPI; 2000-594572/56.		QY	ATTGCTGGCTAGGAGATAACACTGTCTCTTGCCTGATTCGATACCAACGATGTGG	660
XX	Monitoring differential expression of genes in filamentous fungal cells		DB	ATTGCTGGCTAGGAGATAACACTGTCTCTTGCCTGATTCGATACCAACGATGTGG	662
XX	uses fluorescence-labeled nucleic acids isolated from the cells and a		QY	TCAGNATCAATGTTACGACTGGTGGGATCATTTGTTATCGAACTACTCATTGACGGCC	720
XX	substrate of expressed sequence tags.		DB	TCAGNATCAATGTTACGACTGGTGGGATCATTTGTTATCGAACTACTCATTGACGGCC	722
PS	Claim 88; Page 2231-2232; 3161pp; English.		QY	TCGTTATCCACACAGTAAACACAGTCCAGAGGACTTCTGGCCGGGTACAAACAGCCG	780
CC	The present invention describes a method for monitoring differential		DB	TCGTTATCCACACAGTAAACACAGTCCAGAGGACTTCTGGCCGGGTACAAACAGCCG	782
CC	expression of genes in a first filamentous fungal (PF) cell relative to		QY	CAGGCGTGTACTGTATCGCGAGGTGCTCGACGGTGTATCGGCTACACTTGTCCCTACC	840
CC	expression of the same genes in one or more second filamentous fungal		DB	CAGGCGTGTACTGTATCGCGAGGTGCTCGACGGTGTATCGGCTACACTTGTCCCTACC	842
CC	cells. The method uses fluorescence-labeled nucleic acids isolated from		QY	AGAACGTATGAGCGGCTACTGAACTATCCCATTTTACTATCCACTCTCAACGCTTCA	900
CC	the PF cells and a substrate of expressed sequence tags (EST). The ESTs		DB	AGAACGTATGAGCGGCTACTGAACTATCCCATTTTACTATCCACTCTCAACGCTTCA	902
CC	are used in the methods for monitoring differential expression of genes		QY	AGTCAACCTCCCGGACATGGAGACCTCTAACATGATCAACCGTCAATTCGACT	960
CC	in a first filamentous fungal (PF) cell relative to expression of the		DB	AGTCAACCTCCCGGACATGGAGACCTCTAACATGATCAACCGTCAATTCGACT	962
CC	same genes in one or more second filamentous fungal cells. Monitoring the		QY	GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAAACCAACGACCCACGGTTCGCTT	1020
CC	global expression of genes from PF cells allows the production potential		DB	GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAAACCAACGACCCACGGTTCGCTT	1022
CC	of the microorganisms to be improved. New genes may be discovered.		QY	CTTTACCAACGACATAGCCCTCGCCCAAGAAAGTTCGAGATTCATCTCAACGACG	1080
CC	possible functions of unknown open reading frames can be identified and		DB	CTTTACCAACGACATAGCCCTCGCCCAAGAAAGTTCGAGATTCATCTCAACGACG	1082
CC	gene copy number variation and stability can be monitored. The expression		QY	GAATCCCCATCATCTACGCGCGCAAGAACAGCAGCTACGCGCGGAAAGACCCCGCA	1140
CC	of genes can be used to study how PF cells adapt to changes in culture		DB	GAATCCCCATCATCTACGCGCGCAAGAACAGCAGCTACGCGCGGAAAGACCCCGCA	1142
CC	conditions, environmental stress, spore morphogenesis, recombination,		QY	ACCGGAAGCAACCTGGCTCTCTGGGCTACCGGACCGACAGCGAGCTGTACAAGTTAATG	1200
CC	metabolic or catabolic pathway engineering. Using ESTs provides several		DB	ACCGGAAGCAACCTGGCTCTCTGGGCTACCGGACCGACAGCGAGCTGTACAAGTTAATG	1202
CC	advantages over genomic or random cDNA clones including elimination of		QY	CTTCCGCGAAGCAAGCAATCCCGAACTATGCCATTTAGCAAGATACAGGATTCGTGCAACGC	1260
CC	redundancy as one spot on an array equals one gene or open reading frame,		DB		
CC	and organization of the microarrays based on function of the gene				
CC	products to facilitate analysis of the results. AAF07478 to AAF11247				
CC	represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents				
CC	ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from				
CC	Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from				
CC	Trichoderma reesei, which are all specifically claimed in the present				
CC	invention				
XX	Sequence 1914 BP; 484 A; 510 C; 461 G; 459 T; 0 U; 0 Other;				
SQ					
Query Match			98.4%; Score 1707; DB 3; Length 1914;		
Best Local Similarity			100.0%; Pred. No. 0;		
Matches 1707; Conservative			0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	TCATCATCAGCTCTCCGTTCTCTGAAACAATAAACCACAGAGGCAATTTATGATGTCG	60		
DB	3	TCATCATCAGCTCTCCGTTCTCTGAAACAATAAACCACAGAGGCAATTTATGATGTCG	62		
QY	61	CGTGGTGGTCTCTATTCTCTACGGCTTCAGGTTCGCGGACCTGCTTTGGCTGCAACGC	120		
DB	63	CGTGGTGGTCTCTATTCTCTACGGCTTCAGGTTCGCGGACCTGCTTTGGCTGCAACGC	122		

Db 1203 CCTCCGCAACGCAATCCGGAATATGCCATATGCAAGATACAGGATTCGTGACCTACA 1262
Qy 1261 AGAATCTGGCCCATCTACAAAGACGACAAACGATCGCCATCGCAGGCGACAGATGGGT 1320
Db 1263 AGAATCTGGCCCATCTACAAAGACGACAAACGATCGCCATCGCAGGCGACAGATGGGT 1322
Qy 1321 CGCAGATCGTGACTATCTTGTCCAAAGGGTGTCTCGGGTGATTCGATATACCCCTCTCCT 1380
Db 1323 CGCAGATCGTGACTATCTTGTCCAAAGGGTGTCTCGGGTGATTCGATATACCCCTCTCCT 1382
Qy 1381 TGAGTGGTGGGGTTACACAGCGCGCCAGCAATTTGACGAGGTCAATTTGGCTGCACGACCG 1440
Db 1383 TGAGTGGTGGGGTTACACAGCGCGCCAGCAATTTGACGAGGTCAATTTGGCTGCACGACCG 1442
Qy 1441 TGACGGTGGTTCGGATGGAATGTGCTTCTTCCATTCGACGATGGGCTACCTAGGTTAT 1500
Db 1443 TGACGGTGGTTCGGATGGAATGTGCTTCTTCCATTCGACGATGGGCTACCTAGGTTAT 1502
Qy 1501 TGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTGTGAAGGGTGGAGA 1560
Db 1503 TGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTGTGAAGGGTGGAGA 1562
Qy 1561 GTATATGATGGTACTGCTATTCATCTGCAATTCGATTCGACAGTGTGAGTTGATGTACA 1620
Db 1563 GTATATGATGGTACTGCTATTCATCTGCAATTCGATTCGACAGTGTGAGTTGATGTACA 1622
Qy 1621 GTTGGAGTGGTACTGCTGCTGATCCCTTATCTCTTCGATTTGTTTCGAAACCTTAATG 1680
Db 1623 GTTGGAGTGGTACTGCTGCTGATCCCTTATCTCTTCGATTTGTTTCGAAACCTTAATG 1682
Qy 1681 CCAAGCACGCTAGTCTATTATAGGAAA 1707
Db 1683 CCAAGCACGCTAGTCTATTATAGGAAA 1709

RESULT 3
AAQ06786
ID AAQ06786 standard; DNA; 2337 BP.
XX
AC AAQ06786;
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 06-MAR-1991 (first entry)
XX
XX Taka-amylose A gene.
DE
XX
XX Alcohol; ss.
XX
XX Aspergillus oryzae; RIB 40.
XX
XX JP02268685-A.
PN
XX
XX 02-NOV-1990.
XX
XX 07-APR-1989; 89JP-00086787.
PF
XX
XX 07-APR-1989; 89JP-00086787.
PR
XX
XX (JOZO-) JOZO SHIGEN KENKYUSHO.
PA
XX
XX WPI; 1990-372009/50.
XX
XX New gene, vector and transformant of taka:amylose A - isolated from
PT Aspergillus oryzae, used to produce alcohol(s).
PT
XX
PS Claim 1; Fig 1; 6pp; Japanese.
XX
XX The DNA was sequenced from a clone isolated from a A. oryzae RIB 40
CC genomic library. It can be used to express recombinant taka-amylose
CC useful in the prodn. of alcoholic drinks and alcohols. See also AAQ06784
CC and AAQ06785. (Updated on 25-MAR-2003 to correct PA field.) (Updated on

CC 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 2337 BP; 630 A; 600 C; 533 G; 574 T; 0 U; 0 Other;
Query Match 61.3%; Score 1063.8; DB 2; Length 2337;
Best Local Similarity 75.3%; Pred. No. 4.7e-235;
Matches 1700; Conservative 0; Mismatches 7; Indels 552; Gaps 8;
Qy 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTCG 60
Db 68 TCACATCAACCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTCG 127
Qy 61 CGTGGTGGTCTCTAATTTCTGTACGGCTTCAGGTCGGGACCTCTGCTTTGGCTGCAACGC 120
Db 128 CGTGGTGGTCTCTAATTTCTGTACGGCTTCAGGTCGGGACCTCTGCTTTGGCTGCAACGC 187
Qy 121 CTGCGGACTGGCGATCGCAATTCATTTATTTCTTCTCACGGATCGATTTGCAAGACCG 180
Db 188 CTGCGGACTGGCGATCGCAATTCATTTATTTCTTCTCACGGATCGATTTGCAAGACCG 247
Qy 181 ATGGGTTCGACGACTCGGACTTGTAACTACTGCGGATC----- 216
Db 248 ATGGGTTCGACGACTCGGACTTGTAACTACTGCGGATCAGGTGTGTTACTACTAGCTT 307
Qy 217 -----AGAAATACTGTGGTGGAAACATGGCGAGGC 245
Db 308 TCAGAAAGAGGAATGTAAGTCTGATATAGAAATACTGTGGTGGAAACATGGCGAGGC 367
Qy 246 ATCATCGACA----- 255
Db 368 ATCATCGACAAGTAAATTTGCCCTTTATCAAAAAAAGAAAGAAACAGCAAGAAAAA 427
Qy 256 -----AGTTGGACTATATCCAGGGAATGG 280
Db 428 TAAATAAAGAAACTTAGTCTTAACCATCATAGTTGGACTATATCCAGGGAATGG 487
Qy 281 CTTACAGCCATCTGGATCACCCGTTTACAGCCAGCTGCCCGACACACCGCATATGG 340
Db 488 CTTACAGCCATCTGGATCACCCGTTTACAGCCAGCTGCCCGACACACCGCATATGG 547
Qy 341 AGATGCTTACATGGCTACTGGCAGCAGGATAT----- 373
Db 548 AGATGCTTACATGGCTACTGGCAGCAGGATATGTAACGATCGTATTTCTTTGGAAATA 607
Qy 374 -----ATATCTCTCG 383
Db 608 TCTACCTGTCTATTTCTACATCATATGAATACTTGTATGGTGTATAGATATCTCTCG 667
Qy 384 AACGAAACTACGGCACTGCAGATGACTTGAAGGCGCTCTTTCCGCGCTTTCATGAGAG 443
Db 668 AACGAAACTACGGCACTGCAGATGACTTGAAGGCGCTCTCTTCGCGCTTTCATGAGAG 727
Qy 444 GGGATGTATCTTTATGGTGCATGGTGTGCTTACCATAT----- 481
Db 728 GGGATGTATCTTTATGGTGCATGGTGTGCTTACCATATGGTGTGCTTTCGCAACTG 787
Qy 482 -----GGGCTATGATGGAG 495
Db 788 ACTTCGGGATATGGTTCAATTTTCAGTACTGACATGAGTAATATCAGGCGCTATGATGGAG 847
Qy 496 CGGGTAGCTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCCAAGACTACTTCCACC 555
Db 848 CGGGTAGCTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCCAAGACTACTTCCACC 907
Qy 556 CGTTCTGTTTCAATCAAAACTATGAAGATCAGACTCAGGTTGAGGATTTCTGGCTAGGAG 615
Db 908 CGTTCTGTTTCAATCAAAACTATGAAGATCAGACTCAGGTTGAGGATTTCTGGCTAGGAG 967
Qy 616 ATAACACTGTCTCTTCTGCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGAATGGT 675
Db 968 ATAACACTGTCTCTTCTGCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGAATGGT 1027
Qy 676 ACGACTGGGTGGGATCATTTGGTATCGAACTACTCCA----- 711

||||| 1028 ACCACTGGGTGGGATCATTTGGTATCGAACTACTCCAGTAGAATATTTCTCCCTCATCTA 1087
QY 712 -----TTGACGGCTCCGTATCGACACAGTA 737
Db 1088 CAACCTTGGCTGATCGATGATATTACGAAATCAGTTGACGGCTCCGTATCGACACAGTA 1147
QY 738 AATACGCTCCAGAGGACTTCTGGCCGGGTACAACAAGCCGACGGCGTGTACTGTATC 797
Db 1148 AATACGCTCCAGAGGACTTCTGGCCGGGTACAACAAGCCGACGGCGTGTACTGTATC 1207
QY 798 GCGAGGTGCTCGACGGTGTATCGGGCTACACTTGTCCCTTACAGAAACGTCATGACGGC 857
Db 1208 GCGAGGTGCTCGACGGTGTATCGGGCTACACTTGTCCCTTACAGAAACGTCATGACGGC 1267
QY 858 GTACTGAACCTATCCCAT----- 874
Db 1268 GTACTGAACCTATCCCATGTATGGTTCTTCCAAACCATGAGCCTTCTTGAAGTCTCATCTC 1327
QY 875 -----TTACTATCCACTCTCTCAACGCTTCAAGCTCAACCTCCG 912
Db 1328 CTAAAGAAACGGGTAAACACAGTTACTTATCCACTCTCAACGCTTCAAGTCAACCTCCG 1387
QY 913 GCAGCATGGACGACTCTCAACATGATCAACACCGTCAATCCGACTGTCCAGACTCAA 972
Db 1388 GCAGCATGGACGACTCTCAACATGATCAACACCGTCAATCCGACTGTCCAGACTCAA 1447
QY 973 CACTCTGGGCACATTCGTGAGAACCCACCAACCCACGGTTCGGTTC----- 1021
Db 1448 CACTCTGGGCACATTCGTGAGAACCCACCAACCCACGGTTCGGTTCGTAAGTCTTC 1507
QY 1022 -----TTACAC 1027
Db 1508 CTTTTATTTTCCGTTCCCAATTTCCACACAGAACCCACCTCAACAGAGCAAAAGTTACAC 1567
QY 1028 CAACGATAGACCTCGCCAAAGAACGTCGACGATTCATCATCTCAACAGCGGATCCC 1087
Db 1568 CAACGATAGACCTCGCCAAAGAACGTCGACGATTCATCATCTCAACAGCGGATCCC 1627
QY 1088 CATCATCTAGCCGGCCCAAGAACGACTACGCGCGGGAAGAACCCCGCAACCGCGA 1147
Db 1628 CATCATCTAGCCGGCCCAAGAACGACTACGCGCGGGAAGAACCCCGCAACCGCGA 1687
QY 1148 AGCAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAAGTTAATTCGCTCCGC 1207
Db 1688 AGCAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAAGTTAATTCGCTCCGC 1747
QY 1208 GAACGAATCCGGAATATGCCATTAGCAAGATACAGGATTCGTGACCTAC 1260
Db 1748 GAACGAATCCGGAATATGCCATTAGCAAGATACAGGATTCGTGACCTACAGGTAAG 1807
QY 1261 ----- 1260
Db 1808 CACAACCTTAACCATACCTAATGCGCTATCTTCAGAGTATCTGACACAGACATAAT 1867
QY 1261 -----AGAACTGCGCCATCTACAAGACGACACAACGATCGCCATCGCAAGG 1308
Db 1868 CACTGGCAATACAGAATGCGCCATCTACAAGACGACACAACGATCGCCATCGCAAGG 1927
QY 1309 GCACAGATGGTGCAGATCGTGAATCTTGTGTCACAAAGGTCGTCGGGTGATTCGT 1368
Db 1928 GCACAGATGGTGCAGATCGTGAATCTTGTGTCACAAAGGTCGTCGGGTGATTCGT 1987
QY 1369 ATACCTCTCTCTTGTAGTGGTGGGTTACACAGCCGCGCAGCAATTCAGGAGTCAATG 1428
Db 1988 ATACCTCTCTCTTGTAGTGGTGGGTTACACAGCCGCGCAGCAATTCAGGAGTCAATG 2047
QY 1429 GCTGCACGACCGTGAAGGTTGGTTCGGATGGAATATGCTTCTTATGGCAGGTGGGC 1488
Db 2048 GCTGCACGACCGTGAAGGTTGGTTCGGATGGAATATGCTTCTTATGGCAGGTGGGC 2107
QY 1489 TACCTAGGTATTTGATCCGATCGAGAGTTCGACGATGACAGATCTGTAGTACGTCGT 1548
|||||

Db 2108 TACCTAGGTATTTGATCCGACTGAGAACTGCGCAGGTAGCAAGATCTGTAGTACGTCGT 2167
QY 1549 GAAGGTGGAGAGTATATGATGTAAGTCTGTAATCTGCAATCTGGCAATGGACAGTGA 1608
Db 2168 GAAGGTGGAGAGTATATGATGTAAGTCTGTAATCTGCAATCTGGCAATGGACAGTGA 2227
QY 1609 GTTTGATGTACAGTTCGAGTTCGTTACTGCTGTCATCCCTTATACCTTCGATTTGTTT 1668
Db 2228 GTTTGATGTACAGTTCGAGTTCGTTACTGCTGTCATCCCTTATACCTTCGATTTGTTT 2287
QY 1669 CGAACCTTAATGCAAGCACGCTAGTCTATTATAGGAAA 1707
Db 2288 CGAACCTTAATGCAAGCACGCTAGTCTATTATAGGAAA 2326
RESULT 4
ACC44572
ID ACC44572 standard; cDNA; 3285 BP.
XX
AC ACC44572;
DT 02-JUN-2003 (first entry)
XX
Alpha-amylase/glucoamylase fusion nucleotide sequence SEQ ID NO:46.
XX
Self-processing plant; plant; processing enzyme; alpha-amylase; grain;
KW pullulanase; alpha-glucosidase; glucose isomerase; glucoamylase;
KW mesophilic; thermophilic; hyperthermophilic; transgenic plant; starch;
KW maltodextrin; ethanol; fermentation; beverage; enzyme; gene; ss.
XX
Aspergillus shirousami.
OS Synthetic.
XX
FH Key Location/Qualifiers
CDS 1..3285
FT /*tag= a
FT /*partial
FT /*product= "alpha-amylase/glucoamylase fusion protein"
FT /*note= "no start or stop codons given"
XX
WO2003018766-A2.
XX
06-MAR-2003.
XX
27-AUG-2002; 2002WO-US027129.
XX
27-AUG-2001; 2001US-0315281P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Lanahan MB, Basu SS, Batie CJ, Chen W, Craig J, Kinkema M;
XX
WPI: 2003-268420/26.
DR P-PSDB; ABP96630.
XX
Novel polynucleotide encoding hyperthermophilic processing enzymes e.g.
PT alpha-amylase, useful for producing plant to produce food products having
PT improved taste or fermentable substrates for ethanol.
XX
Claim 1; Page 107-109; 158pp; English.
XX
The present invention describes polynucleotides which encode processing
CC enzymes (e.g. alpha-amylase, pullulanase, alpha-glucosidase, glucose
CC isomerase, or glucoamylase) that are optimised for expression in plants.
CC The polynucleotides encode mesophilic, thermophilic or hyperthermophilic
CC processing enzymes, which are activated under suitable conditions to act
CC upon the desired substrate. Also described are self-processing transgenic
CC plants and plant parts, e.g. grain, which express one or more of these
CC enzymes and have an altered composition that facilitates plant and grain
CC processing. Also described is a method (M) for converting starch to
CC starch-derived products in a transformed plant part (TPP), by activating
CC the starch processing enzyme contained in it. Transgenic grain is useful
CC for preparing maltodextrin. A transformed plant (TP) can be used to

PT Producing heterologous biological substance comprises culturing mutant of
PT wild-type *Aspergillus niger* strain in medium suitable for producing
PT heterologous biological substance and recovering heterologous biological
PT substance.

XX Example 10; SEQ ID NO 17; 58pp; English.

XX The present invention relates to a method of producing heterologous
CC biological substance. The method involves culturing mutant of wild-type
CC *Aspergillus niger* strain in medium suitable for producing heterologous
CC biological substance, where mutant strain comprises first nucleotide
CC sequence encoding heterologous biological substance and second nucleotide
CC sequences comprising modification of glucoamylase (glua) and recovering
CC heterologous biological substance. The present sequence is the
CC *Aspergillus niger* neutral alpha-amylase B (amyB) gene.

XX Sequence 2443 BP; 669 A; 620 C; 558 G; 595 T; 0 U; 1 Other;

Query Match 56.5%; Score 979; DB 13; Length 2443;
Best Local Similarity 74.4%; Pred. No. 1.7e-215;
Matches 1609; Conservative 0; Mismatches 10; Indels 544; Gaps 8;
1 TCACATCAAGCTCCCTCTCTGAACAATAAAACCCACAGAGGCAATTTATGATGTCG 60
151 TCACATCAAGCTCCCTCTCTGAACAATAAAACCCACAGAGGCAATTTATGATGTCG 210
61 CGTGGTGTCTTATTTCTGACGGCTTCAGTCCGGCACCCTGTTGGTGCACACGC 120
211 CGTGGTGTCTTATTTCTGACGGCTTCAGTCCGGCACCCTGTTGGTGCACACGC 270
121 CTGGGACTGGCGATCGCAATCATTATTTCTCTCAGGATCGATTTGCAAGACGG 180
271 CTGGGACTGGCGATCGCAATCATTATTTCTCTCAGGATCGATTTGCAAGACGG 330
181 ATGGGTGCGAGCTGGGACTTGAATCTGCGATC----- 216
331 ATGGGTGCGAGCTGGGACTTGAATCTGCGATCAGGTGTTGTTACTACTAGCTT 390
217 -----AGAAATACTGTGGTGGAAATCGCAGGGC 245
391 TCAGAAAGAGGAATGATAAAGTAACTGATATAGAAATACTGTGTGGAACATGCGAGGC 450
246 ATCATCGACA----- 255
451 ATCATCGACAAGGTAAATTGCCCTTTATCAAAAAAAGAGGAAAAAGCAGAAAGAAA 510
256 -----AGTTGGACTATATCCAGGGAATGG 280
511 TAAAAATAAAGAACTAGTCTTAACCATCAATATGTTGGACTATATCCAGGGAATGG 570
281 CTTTACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCGACACACCCGATATGG 340
571 CTTTACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCGACACACCCGATATGG 630
341 AGATGCTTACCATGGCTTCTGGCAGCAGGATAT----- 373
631 AGATGCTTACCATGGCTTCTGGCAGCAGGATATGTAAGTCTTTTAAATATCTACC 690
374 -----ATACTCTCTGACGAAA 391
691 TGTCACTTTTTCATCAATATGAATCACTTGTATGTTTATAGATATCTCTGAAACGAAA 750
392 CTACGGCACTGCAGATGACTTGAAGCGCTCTCTTGGGCCCTTTCATGAGAGGGGATGTA 451
751 CTACGGCACTGCAGATGACTTGAAGCGCTCTCTTGGGCCCTTTCATGAGAGGGGATGTA 810
452 TCTTATGGTGCATGTGGTGGTCTAACCATAT----- 481
811 TCTTATGGTGCATGTGGTGGTCTAACCATATGGTGGTGGTGGTGGTGGTGGTGGTGG 870
482 -----GGGCTATGATGAGCGGGTAGC 503
871 GATATGTTTCATTTTCACTGACTGACAAATAGTAAATATATCAGGGCTATGATGGAGGGGTAGC 930

QY 504 TCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCTCCAGACTACTTCCACCGCTTCTGT 563
DB 931 TCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCTCCAGACTACTTCCACCGCTTCTGT 990
QY 564 TTCATTCAAACTATGAAGATCAGACTCAGGTTTGAGGATTTGCTGGCTAGGAGATAACACT 623
DB 991 TTCATTCAAACTATGAAGATCAGACTCAGGTTTGAGGATTTGCTGGCTAGGAGATAACACT 1050
QY 624 GTCTCCTTCCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGAATGGTACGACTGG 683
DB 1051 GTCTCCTTCCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGAATGGTACGACTGG 1110
QY 684 GTGGGATCATTTGGTATCGAACTACTCCA----- 711
DB 1111 GTGGGATCATTTGGTATCGAACTACTCCAAGTAAATATTTCTCCCTCATTTCTACAACCTGG 1170
QY 712 -----TTGACGGCCTCCGTATCGACACAGTAAACACCT 745
DB 1171 CTGATCGATGATACATTACGAATCAGTTGACGGCCTCCGTATCGACACAGTAAACACCT 1230
QY 746 CCAGAGGACTTTCTGGCCCGGGGTACAAACAAAGCCGAGCGGTGTACTGTATCGCGAGGT 805
DB 1231 CCAGAGGACTTTCTGGCCCGGGGTACAAACAAAGCCGAGCGGTGTACTGTATCGCGAGGT 1290
QY 806 GCTCGAGCGTGTATCGGCGCTACACTTGTCTCCTACCAAGACGTCATGACGCGGTACTGAA 865
DB 1291 GCTCGAGCGTGTATCGGCGCTACACTTGTCTCCTACCAAGACGTCATGACGCGGTACTGAA 1350
QY 866 CTATCCCAT----- 874
DB 1351 CTATCCCATGTATGGTTCCTCAACCATGAGCCTCTTGCAGTCTCATCTCTCAACGAA 1410
QY 875 -----TTACTATCCACTCTCAAGCCTTCAAGTCAACCTCCGCGAGCATG 920
DB 1411 ACGGCTAAACAGTTTACTTATCCACTCTCAAGCCTTCAAGTCAACCTCCGCGAGCATG 1470
QY 921 GACGACTCTCAACATGATCAACCGGTCAAAATCCGACTGTCCAGACTCAACACTCCTCTG 980
DB 1471 GACGACTCTCAACATGATCAACCGGTCAAAATCCGACTGTCCAGACTCAACACTCCTCTG 1530
QY 981 GGCACATTCGTGAGAACCAACGACACACCGGTTCGCTTC----- 1021
DB 1531 GGCACATTCGTGAGAACCAACGACACACCGGTTCGCTTCGTAAGTCTTCCTCTTTTATT 1590
QY 1022 -----TTACACCAACGACA 1035
DB 1591 TTCCGTTTCCCAATTTCCACACAGAACCCACCTTAAACAGACAAAGTTACACCAACGACA 1650
QY 1036 TAGCCCTCGCCCAAGAACGTCGACGATTCATCATCTCAACGAGAAATCCCATCATCT 1095
DB 1651 TAGCCCTCGCCCAAGAACGTCGACGATTCATCATCTCAACGAGAAATCCCATCATCT 1710
QY 1096 AGCCGGCCCAAGAACGACTAGCCCGCGGAAACGACCCCGGAAACCGGAGAACACCT 1155
DB 1711 AGCCGGCCCAAGAACGACTAGCCCGCGGAAACGACCCCGGAAACCGGAGAACACCT 1770
QY 1156 GGCTCTCGGGCTACCCGACGAGCGAGCTGTACAAAGTTAAATTTGCTCCGCGAACGCAA 1215
DB 1771 GGCTCTCGGGCTACCCGACGAGCGAGCTGTACAAAGTTAAATTTGCTCCGCGAACGCAA 1830
QY 1216 TCCGGAATATGCCATTAGCAAGATATAGGATTCGTGACCTACA----- 1260
DB 1831 TCCGGAATATGCCATTAGCAAGATATAGGATTCGTGACCTACAAGGTAAGCACAACCT 1890
QY 1261 ----- 1260
DB 1891 CTAAGCATACCTTAATGGCTATCTTTCAGAGTATCTGACACAAAGAGACTAATCACTGGCA 1950
QY 1261 ----AGAACTGGCCCATCTACAAAGACGACACAAAGTCCCATCGCAAGGCGACAGAT 1316
DB 1951 ATACAGAACTGGCCCATCTACAAAGACGACACAAAGTCCCATCGCAAGGCGACAGAT 2010

QY 1317 GGGTCGACAGTCGTGACTATCTTGTCCAAACAGGGTGTCTCGGGTGATTCGTATACCCCTC 1376
Db |||||
2011 GGGTCGACAGTCGTGACTATCTTGTCCAAACAGGGTGTCTCGGGTGATTCGTATACCCCTC 2070
QY 1377 TCCTTGAGTGTGCGGGTTACACAGCCGCGCAGCAATTGACGAGGTCAATTGGCTGCAGC 1436
Db |||||
2071 TCCTTGAGTGTGCGGGTTACACAGCCGCGCAGCAATTGACGAGGTCAATTGGCTGCAGC 2130
QY 1437 ACCGTGACGGTTGGTTCGGATGGAATGTGCTGTCTTCTATGCGAGGTGGGTACTAGG 1496
Db |||||
2131 ACCGTGACGGTTGGTTCGGATGGAATGTGCTGTCTTCTATGCGAGGTGGGTACTAGG 2190
QY 1497 GTATTGTATCCGACTGAGAGTGTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 1556
Db |||||
2191 GTATTGTATCCGACTGAGAGTGTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 2250
QY 1557 GAGAGTATATGATGTACTGTCTATTCAATCTGGCATTGACAGTAGTGAGTTGATG 1616
Db |||||
2251 GAGAGTATATGATGTACTGTCTATTCAATCTGGCATTGACAGTAGTGAGTTGATG 2310
QY 1617 TAC 1619
Db ||
2311 GAC 2313

RESULT 6
ADT89631
ID ADT89631 standard; DNA; 2520 BP.
XX
AC ADT89631;
XX
DT 16-DEC-2004 (first entry)
XX
DE Aspergillus niger neutral alpha-amylase A (amyA) gene.
XX
KW Glucoamylase; glaB; amyA; alpha-amylase A; gene; ds.
XX
OS Aspergillus niger.
XX
FH Location/Qualifiers
FT 201..2244
FT /*tag= a
FT /product= "Neutral alpha-amylase A (amyA) protein"
FT /note= "CDS has many translational exceptions which
FT alters the reading frame"
XX
PN US2004191864-A1.
XX
XX 30-SEP-2004.
XX
XX 31-MAR-2004; 2004US-00815495.
XX
XX 31-MAR-2003; 2003US-0459902P.
XX
XX (NOVO) NOVOZYMES BIOTECH INC.
XX
XX Connelly M, Brody H;
XX
XX WPI; 2004-708545/69.
XX
XX P-PSDB; ADT89632.
XX
XX Producing heterologous biological substance comprises culturing mutant of
XX wild-type Aspergillus niger strain in medium suitable for producing
XX heterologous biological substance and recovering heterologous biological
XX substance.
XX
XX Example 11; SEQ ID NO 21; 58pp; English.
XX
XX The present invention relates to a method of producing heterologous
XX biological substance. The method involves culturing mutant of wild-type
XX Aspergillus niger strain in medium suitable for producing heterologous
XX biological substance, where mutant strain comprises first nucleotide
XX sequence encoding heterologous biological substance and second nucleotide

CC sequences comprising modification of glucoamylase (glaA) and recovering
CC heterologous biological substance. The present sequence is the
CC Aspergillus niger neutral alpha-amylase A (amyA) gene.
XX
SQ Sequence 2520 BP; 682 A; 645 C; 581 G; 611 T; 0 U; 1 Other;
Query Match 52.4%; Score 908.2; DB 13; Length 2520;
Best Local Similarity 73.6%; Pred. No. 3.5e-199;
Matches 1537; Conservative 0; Mismatches 8; Indels 544; Gaps 8;
QY 1 TCACATCAAGCTCTCCCTCTCTGAAACAATAAACCCACAGAGGCAATTTATGATGGTGG 60
Db 151 TCACATCAAGCTCTCCCTCTCTGAAACAATAAACCCACAGAGGCAATTTATGATGGTGG 210
QY 61 CGTGTGTGTCTCTATTTCTGTACGGCTTCAGGTGCGCGCACCTGCTTTGGCTGCAACGC 120
Db 211 CGTGTGTGTCTCTATTTCTGTACGGCTTCAGGTGCGCGCACCTGCTTTGGCTGCAACGC 270
QY 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCTTCTCAAGATCGATTTGCAAGACGG 180
Db 271 CTGCGGACTGGCGATCGCAATCCATTTATTTCTTCTCAAGATCGATTTGCAAGACGG 330
QY 181 ATGGTTCGACGACTCGGACTTGTAACTACGCGATC----- 216
Db 331 ATGGTTCGACGACTCGGACTTGTAACTACGCGATCAGGTGTGTGTACCTACTAGCTTT 390
QY 217 -----AGAAATACTGTGTGGGAACATGCGCAGGC 245
Db 391 TCAGAAAGAGGAAATGTAACGACTTGATATAGAAATACTGTGTGGGAACATGCGCAGGC 450
QY 246 ATCATCGACA----- 255
Db 451 ATCATCGACAAGGTAAATTGCCCCCTTTATCAAAAAAAGAGGAAAGAGCAAGAAAA 510
QY 256 -----AGTTGCACTATATCCAGGGAATGG 280
Db 511 TAAAAATAAAAGAACTCTAGTCTTAACCATCATCATAGTTGGACTATATCCAGGGAATGG 570
QY 281 CTTACAGCCATCTCGATCACCCCGTTACAGCCAGCTGCCACAGCACCGCATATGCG 340
Db 571 CTTACAGCCATCTCGATCACCCCGTTACAGCCAGCTGCCACAGCACCGCATATGCG 630
QY 341 AGATCCCTACCATGGCTACTTGGCAGCAGGAT----- 373
Db 631 AGATCCCTACCATGGCTACTTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC 690
QY 374 -----ATACTCTGAAACGAAAA 391
Db 691 TGTCATCTTTTACATCAATATGAACCTAACTTGATGGTTTATAGATACTCTCTGAACGAAAA 750
QY 392 CTACGGCACTGCAGATGACTTGAAGCGCTCTCTCGGCCCTTCATGAGAGGGGATGTA 451
Db 751 CTACGGCACTGCAGATGACTTGAAGCGCTCTCTCGGCCCTTCATGAGAGGGGATGTA 810
QY 452 TCTTATGGTCGATGTGTGTGCTTAACCAT----- 481
Db 811 TCTTATGGTCGATGTGTGTGCTTAACCATATGGTTCGTGTCTCTTTCGCAACTGACTTTCGCG 870
QY 482 -----GGGCTATGATGGAGGGGTAGC 503
Db 871 GATATGGTTTCACTTTTCAGTACTGACAATGAGTAATATCAGGGCTATGATGGAGGGGTAGC 930
QY 504 TCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCTCCAGACTACTTCCACCGCTTCTGT 563
Db 931 TCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCTCCAGACTACTTCCACCGCTTCTGT 990
QY 564 TTCATTCAAACTATCAAGATCAGACTCAGGTGTGAGGATTTGCTGCTAGGAGATAACACT 623
Db 991 TTCATTCAAACTATCAAGATCAGACTCAGGTGTGAGGATTTGCTGCTAGGAGATAACACT 1050
QY 624 GTCTCTTCCCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGATGGTACGACTGG 683
Db 1051 GTCTCTTCCCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGATGGTACGACTGG 1110


```
QY 684 GTGGATCATTTGTTATCGAAGTACTCA-----711
Db 1111 GTGGATCATTTGTTATCGAAGTACTCAAGTAAGATATTTCTCCCTCATTTCAACATTGG 1170
712 -----TTGACGGCTCCGTTATCGACACAGTAAACACAGT 745
Db 1171 CTGATCGATGATATTACGAATCAGTTGACGGCTCCGTTATCGACACAGTAAACACAGT 1230
QY 746 CAGAAGGACTTCTGGCCCGGTACAAAGGCGGAGGCTGTATCTGTATCGGAGGT 805
Db 1231 CAGAAGGACTTCTGGCCCGGTACAAAGGCGGAGGCTGTATCTGTATCGGAGGT 1290
QY 806 GCTCGAGGATCGCGCTACACTTGTCCCTACCAAGCTCATGGACGGCTACTGAA 865
Db 1291 GCTCGAGGATCGCGCTACACTTGTCCCTACCAAGCTCATGGACGGCTACTGAA 1350
QY 866 CTATCCAT-----874
Db 1351 CTATCCATGTTGTTCTCCCAACATGAGCTTCTTGCAAGTCTCATCTCTTAAGAA 1410
QY 875 -----TTACTATCCACTCTCAAGCCCTTCAAGTCAACCTCCGGCAGCATG 920
Db 1411 ACGGCTAAACACGTTACTTATCACTCTCAAGCCCTTCAAGTCAACCTCCGGCAGCATG 1470
QY 921 GAGGACTCTAACAATGATCAACCGTCAAAATCGACTGTCCAGACTCAACACTCTTG 980
Db 1471 GAGGACTCTAACAATGATCAACCGTCAAAATCGACTGTCCAGACTCAACACTCTTG 1530
QY 981 GGCACATTGTCGAGAACACGACCAACCCACGGTTCGCTTC-----1021
Db 1531 GGCACATTGTCGAGAACACGACCAACCCACGGTTCGCTTCGTTGAAGTCTTCCTTTTATT 1590
QY 1022 -----TTACACCAACGACA 1035
Db 1591 TTCGTTTCCCAATTTCCACACAGAACCCCACTTAACAGAGCAAGTTACACCAACGACA 1650
QY 1036 TAGCCCTCGCAAGAGCTCGAGATTCATCTCTCAAGCGGAATCCCATCATCT 1095
Db 1651 TAGCCCTCGCAAGAGCTCGAGATTCATCTCTCAAGCGGAATCCCATCATCT 1710
QY 1096 ACGCGGCGCAAGACAGCTACCGCGGGAACGACCCCGGAACCGCGGAGCAACCT 1155
Db 1711 ACGCGGCGCAAGACAGCTACCGCGGGAACGACCCCGGAACCGCGGAGCAACCT 1770
QY 1156 GGCTCTCGGCTACCGACCGAGAGCTGTACAAAGTTAAATGCTTCGCGCAACGCAA 1215
Db 1771 GGCTCTCGGCTACCGACCGAGAGCTGTACAAAGTTAAATGCTTCGCGCAACGCAA 1830
QY 1216 TCCGGAATATGCTATAGCAAGATACAGGATTCGTGACCTACA-----1260
Db 1831 TCCGGAATATGCTATAGCAAGATACAGGATTCGTGACCTACAAGGTAAGCAACCT 1890
QY 1261 -----1260
Db 1891 CTAAGCATACCTTAATGGCCATCTTTCAGAGTATCTGACACAAGAGACTAATCACTGGCA 1950
QY 1261 ----AGAACTGGCCATCTACAAGAGGACACAACGATCCCATCGCAAGGGCACAGAT 1316
Db 1951 ATACAGACTGGCCATCTACAAGAGGACACAACGATCCCATCGCAAGGGCACAGAT 2010
QY 1317 GGCTCGCAGATCGTACTATCTTGTCCAAAGGGGCTTCGGGGTGAATCTGTATACCCCTC 1376
Db 2011 GGCTCGCAGATCGTACTATCTTGTCCAAAGGGGCTTCGGGGTGAATCTGTATACCCCTC 2070
QY 1377 TCCTTAGTGGTGGGGTTACACAGCGGCGCAGCAATTCGAGAGTTCATTTGGCTGCAGC 1436
Db 2071 TCCTTAGTGGTGGGGTTACACAGCGGCGCAGCAATTCGAGAGTTCATTTGGCTGCAGC 2130
QY 1437 ACCGTGACGGTTCGTTTCGAGTGAATATGTCCTGTTCTTATGTCAGTGGCTACCTAGG 1496
Db 2131 ACCGTGACGGTTCGTTTCGAGTGAATATGTCCTGTTCTTATGTCAGTGGCTACCTAGG 2190
```

```
QY 1497 GTATTCTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCT 1545
Db 2191 GTATTCTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCT 2239

RESULT 7
AAF11252
ID AAF11252 standard; cDNA; 1389 BP.
XX AAF11252;
XX 13-MAR-2001 (first entry)
XX Aspergillus niger EST SEQ ID NO:3775.
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX Aspergillus niger.
XX OS
XX PN WO200056762-A2.
XX XX
XX PD 28-SEP-2000.
XX XX
XX PF 22-MAR-2000; 2000WO-US007781.
XX PR 22-MAR-1999; 99US-00273623.
XX PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX PA (NOVO ) NOVO NORDISK AS.
XX PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX WPI; 2000-594572/56.
XX
XX PT Monitoring differential expression of genes in filamentous fungal cells
XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a
XX PT substrate of expressed sequence tags.
XX PS Claim 87; Page 1694-1695; 3161pp; English.
XX
CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring the
CC global expression of genes from FF cells allows the production potential
CC of the microorganisms to be improved. New genes may be discovered,
CC possible functions of unknown open reading frames can be identified and
CC gene copy number variation and stability can be monitored. The expression
CC of genes can be used to study how FF cells adapt to changes in culture
CC conditions, environmental stress, spore morphogenesis, recombination,
CC metabolic or catabolic pathway engineering. Using ESTs provides several
CC advantages over genomic or random cDNA clones including elimination of
CC redundancy as one spot on an array equals one gene or open reading frame,
CC and organisation of the microarrays based on function of the gene
CC products to facilitate analysis of the results. AAF07478 to AAF11247
CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from
CC Trichoderma reesei, which are all specifically claimed in the present
CC invention
XX
SQ Sequence 1389 BP; 354 A; 384 C; 325 G; 326 T; 0 U; 0 Other;
Query Match 46.5%; Score 806; DB 3; Length 1389;
Best Local Similarity 100.0%; Pred. No. 1.1e-175;
```

		Matches	806;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	902	GTCAACCTCCGGCAGCATGACGACCTCTACAACATGATCAACACCGTCAAAATCCGACTG	961								
Db	1	GTCAACCTCCGGCAGCATGAGACCTCTACAACATGATCAACACCGTCAAAATCCGACTG	60								
Qy	962	TCAGACTCAACACTCTCGGGCACATTCGTGAGAACCAACGACCCACGGTTCGCTTC	1021								
Db	61	TCAGACTCAACACTCTCGGGCACATTCGTGAGAACCAACGACCCACGGTTCGCTTC	120								
Qy	1022	TTACACCAACGATAGCCCTCGCCAGAACGTGCGAGATTCATCATCTCTCAACGACGG	1081								
Db	121	TTACACCAACGATAGCCCTCGCCAGAACGTGCGAGATTCATCATCTCTCAACGACGG	180								
Qy	1082	AATCCCCATCATCTAGCCGGCCCAAGAACAGACACTACGCCGCGGAAGACCCCGCGAA	1141								
Db	181	AATCCCCATCATCTAGCCGGCCCAAGAACAGACACTACGCCGCGGAAGACCCCGCGAA	240								
Qy	1142	CCGGAAGCAACTGGCTCTCGGGCTACCCGACCGACGAGCTGTACAAGTTAATTGC	1201								
Db	241	CCGGAAGCAACTGGCTCTCGGGCTACCCGACCGACGAGCTGTACAAGTTAATTGC	300								
Qy	1202	CTCCGGGAACGGAATCCGGAATATGCCATTAGCAAAAGATACAGATTTCGTGACCTACAA	1261								
Db	301	CTCCGGGAACGGAATCCGGAATATGCCATTAGCAAAAGATACAGATTTCGTGACCTACAA	360								
Qy	1262	GAACTGGCCATCTACAAAGACACACACGATCGCCATCGCAAGGGCACAGATGGGTC	1321								
Db	361	GAACTGGCCATCTACAAAGACACACACGATCGCCATCGCAAGGGCACAGATGGGTC	420								
Qy	1322	GCAGATCGTGACTATCTTGTCCAAAGGGTGCTTCGGGTGATTCGTATACCCCTCTCCTT	1381								
Db	421	GCAGATCGTGACTATCTTGTCCAAAGGGTGCTTCGGGTGATTCGTATACCCCTCTCCTT	480								
Qy	1382	GAGTGTGCGGGTTACAGACCGCGCCAGCAATTGACGAGGTTCATTTGGTCACGACCGT	1441								
Db	481	GAGTGTGCGGGTTACAGACCGCGCCAGCAATTGACGAGGTTCATTTGGTCACGACCGT	540								
Qy	1442	GACGTTGCTCGGATGGAATGTGCTCTTCTATGGCAGGTGGCTACCTAGGGTATT	1501								
Db	541	GACGTTGCTCGGATGGAATGTGCTCTTCTATGGCAGGTGGCTACCTAGGGTATT	600								
Qy	1502	GTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTCTAGTAGCTCTGTAAGGGTGGAGAG	1561								
Db	601	GTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTCTAGTAGCTCTGTAAGGGTGGAGAG	660								
Qy	1562	TATATAGTGTACTGCTATTCAATCTGGCAATGGACAGTGAGTTTGATGTACAG	1621								
Db	661	TATATAGTGTACTGCTATTCAATCTGGCAATGGACAGTGAGTTTGATGTACAG	720								
Qy	1622	TTGGAGTCGTACTGCTGTCATCCCTTATCTCTTCGATTGTTTTTCGAACCCCTAATGC	1681								
Db	721	TTGGAGTCGTACTGCTGTCATCCCTTATCTCTTCGATTGTTTTTCGAACCCCTAATGC	780								
Qy	1682	CAAGCAGCTAGTCTATTATTAGAAA	1707								
Db	781	CAAGCAGCTAGTCTATTATTAGAAA	806								

RESULT 8
ABQ80346
ID ABQ80346 standard; cDNA; 1488 BP.
XX ABQ80346;
AC ABQ80346;
DT 11-AUG-2003 (first entry)
XX A. fumigatus AFAA1 cDNA.
XX Gene; enzyme; tannase; cellulase; glucose oxidase; glucoamylase; phytase;
KW beta-galactosidase; invertase; lipase; alpha-amylase; laccase;
KW polygalacturonase; xylanase; gallate ester linkage; detergent; cellulose;
KW glucose; oxygen; myo-inositol phosphate; lactose; tea leaf; sucrose;

glyceride; starch; maltodextrin; oxidated phenolic compound;
polygalacturonic acid chain; xylan; xylo-oligomer; food; feed; beverage;
textile; tea liquor; cleaning ability; ss.
OS Aspergillus fumigatus.
XX
Key Location/Qualifiers
CDS 1..1488
/*tag= a
/product= "Alpha amylase 1"
WO2003012071-A2.
13-FEB-2003.
XX
05-AUG-2002; 2002WO-US024842.
XX
03-AUG-2001; 2001US-0309870P.
XX (ELIT-) ELITRA PHARM INC.
XX
Jiang B, Storms R, Roemer T, Bussey H;
WPI; 2003-332729/31.
P-PSDB; ABB80177.
XX
Novel isolated Aspergillus fumigatus polypeptide, useful in various industries such as those involved in the making of food and feed, beverages, textiles and detergents.
XX
Claim 2; Page 132-34; 169pp; English.
XX
The sequences given in ABQ80319-66 encode enzymatic proteins derived from A. fumigatus. The resulting proteins display the catalytic activity of an enzyme such as tannase, cellulase, glucose oxidase, glucoamylase, phytase, beta-galactosidase, invertase, lipase, alpha-amylase, laccase, polygalacturonase or xylanase. Compositions comprising the tannase are useful for modulating the amount of compounds that comprise a gallate ester linkage in a composition. Compositions comprising cellulase are useful for modulating the amount of cellulose in a composition. Compositions comprising glucose oxidase are useful for modulating the amount of glucose or oxygen in a composition. Compositions comprising phytase are useful for modulating the amount of myo-inositol phosphates in a composition. Compositions comprising beta-galactosidase are useful for modulating the amount of lactose in a composition. Compositions comprising sucrose or invertase are useful for modulating the amount of sucrose in a composition. Compositions comprising lipase are useful for modulating the amount of glyceride in a composition. Compositions comprising alpha-amylases are useful for modulating the amount of starches or maltodextrins in a composition. Compositions comprising laccase are useful for modulating the amount of oxidated phenolic compounds in a composition. Compositions comprising polygalacturonases are useful for modulating the amount of high or low molecular weight polygalacturonic acid chains in a composition. Compositions comprising xylanases are useful for modulating the amount of xylan or xylo-oligomers in a composition. The A. fumigatus proteins and corresponding DNA's are useful in various industries such as those involved in the making of food and feed, beverages, textiles and detergents. The DNA's are useful to express recombinant enzymes for characterization, modification or industrial uses, to compare with the nucleotide sequence of A.fumigatus to identify duplicated genes of paralogs having the same or similar biochemical activity and/or function, to compare with nucleic acid sequence of other related or distant fungal organisms to identify potential orthologous enzyme genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, and to raise anti-protein antibodies. The polypeptide having tannase activity increases the yield of tea liquor from tea leaves, improves the colour, flavour and health benefits of tea products, particularly an instant tea product. The polypeptide having cellulase activity enhances cleaning ability of detergent compositions

Sequence 1488 BP; 364 A; 410 C; 358 G; 356 T; 0 U; 0 Other;

Query Match 45.3%; Score 786; DB 10; Length 1488;			
Best Local Similarity 71.4%; Pred. No. 4.5e-171;			
Matches 1035; Conservative 0; Mismatches 415; Indels 0; Gaps 0;			
QY	94	TCGCGGACCCCTGCTTTGGCTGCAACGCCCTGCGACTGCGGATCGCAATCCATTATTTC 153	
Db	38	TTGCTGGACAGCTTTGGCGGCCACTCCAGCGCAATGGCGATCCAGTCCATCTATTCT 97	
QY	154	TTCTCAAGGATCGATTGGCAAGACCGATGGGTGACGAGCTGGGACTTTGTAATACCTGCGG 213	
Db	98	TACTCAGGATCGATTTCGAAGGACTGACGGTTCCCAACGGCCTCGTGTGATACCTAGTG 157	
QY	214	ATCAGAAATACCTGTGTGGAACATGCGAGGCGCATATCGACAGTTGGACTATATCCAGS 273	
Db	158	CTCGAGAGTATTGTGTGGTACATGCAAGGCAATTATTGAACAACTCGATTATACATAAG 217	
QY	274	GAATGGGCTTACAGGCATCTGGATCAACCCCGTTACAGCCAGCTGCCCCAGACACCG 333	
Db	218	GCATGGGCTTTACAGCAATTTGATTAACCCAGTCAACAGCAACTTCCCAGGATACGT 277	
QY	334	CATATCGAGATGCTTACCATGGCTACTGGCAGCAGGATATATCTCTTGAACGAAACT 393	
Db	278	CGGAAGGCACTGCATACCAACGGTACTGGCAGCAGACATTTATTCCGTCAATTTCCAAT 337	
QY	394	ACGGCACTGCAGATGACTTGAAGCGCTCTCTTGGCGCCTTTCATGAGAGGGGATGTATC 453	
Db	338	ATGGGACCGTGACGACCTGAAAGCCCTGGCGTCAGCTCTTCATGACAGGGGATGTATC 397	
QY	454	TTATGCTCGATGTGGTTGCTTAACATATGGGCTATGATGAGCGGCTAGCTCACTCGATT 513	
Db	398	TCATGCTCGATGTGGTTGCTTAACATATGGGCTATGATGAGCGGCTAGCTCACTCGATT 457	
QY	514	ACAGTGTGTTAAACCGTTCACTTCCCAAGACTTTCACCCCGTTCTGTTCATTCATAA 573	
Db	458	ATAGTGTCTTCAACCCCTTCAACTCCAGACCTCTTCCACCCATTTGTGCTTCATCAGCA 517	
QY	574	ACTATGAAGATCAGATCAGTTGAGGATGCTGGCTAGGATATACACTGTCTCTTCG 633	
Db	518	ATTACGATATCAGACAGACGTTGGAATTTGCTGGTTGGGGGCAATTCGGTTCTTTAC 577	
QY	634	CTGATCTCGATACACCAAGATGTGTCGAAGATGAATGTCAGCTGGGTGGGATCAT 693	
Db	578	CGATCTTGTACTACAAATCCGATGTTCAAAAGATTTGGTACATTTGGTGAATCTTT 637	
QY	694	TGATATCGAATCTACTCCATGACGGCTCCGATTCGACACAGTAAACACGTCGAGAAG 753	
Db	638	TAGTGTCTAATTATTCATCGACGGCTACGAATTCACACTGTGAAGCAGCTCCAGAGCG 697	
QY	754	ACTTTCGGCCGGGTACAAACAGCCGAGCGGTGACTGTATCGGGGAGGTGCTCGACG 813	
Db	698	ATTTCTGGCCGGATTCATATGACCGCGCAGCGCTTACTGTATCGGGGAGGTATTTGATG 757	
QY	814	GTGATCGGCTACACTTGTCCCTACAGAACGTCATGACGCGGTACTGAACCTATCCCA 873	
Db	758	GAGATCCAGGTACACTTGTCCCTACCAAGAGTCTTGAACGGGTGCTGAATATCCCTA 817	
QY	874	TTTACTATCCACTCTCAACGCTTCAAGTCAACCTCCGSCAGCATGGACGCTCTTACA 933	
Db	818	TATATTATCGCTTCTGAAGCTTCCAATCCACAGTGGCAGTATGAGCAGCTCTACG 877	
QY	934	ACATGATCAACACCGTCAAAATCCGACTGTTCAGACTCAACACTCTCTGGGCACATTCGTG 993	
Db	878	ATATGATCAACACCGTCAAGTCGAGTCGCGCCGATTCGACGCTGTTGGGCACCTTTGTG 937	
QY	994	AGAACACGACACCCACGGTTCGCTTCTTACACCAACGACATAGCCCTCGCCCAAGACG 1053	
Db	938	AAATCAGCATATCTCCCGGTTTCGATTCGATATACCAAGGACATGGCCCTCGCTAAACACG 997	
QY	1054	TCGAGCATTCATCATCTCTCAACGAGGAAATCCCATCTCATCTACGCGCGGCAAGAACAGC 1113	
Db	998	CCGAGCATTCATCATCTCTTCGATGGAAATCCCATCATCTACGCGCGGCGCAGAGCAAC 1057	
QY	1114	ACTACCGCGCGGAAACGACCCCGCGAACCGCGAAGCAACCTTGGCTCTCGGGCTACCCGA 1173	
Db	1058	ATTACAGCGCGCGCGGATCCCGNAACCGTAGGCGACTCTGGCTATCCGGCTACTCGA 1117	
QY	1174	CCGACAGCGAGCTGTCAAAGTTAATTTGCTCCGCGAACCGCAATCCGGAATATGCAATTA 1233	
Db	1118	CAACGAGTGACCTATACAGCTCATCGCAGACAGMAACGCCATCCGGAGTCAGGCCATTA 1177	
QY	1234	GCNAAGATACAGATTCGTGACTACAGAACTGGCCCATCTTCAAGAGCAGACACACGA 1293	
Db	1178	GCAAGGACCCGGGATACGTGACTTATAAGAAATAACCCCATCTCAAAAGATACCTCCACCA 1237	
QY	1294	TCGCCATGGCAAGGCGACAGATGGGTCCGAGATCGTGACTATCTTGTCCAACAAGGGTG 1353	
Db	1238	TGCCATGGCAAAAGGCTCCGACGAGCGCAGATCATCCGCTCTCTCGAACCTCGGTG 1297	
QY	1354	CTTCGGGTGATTCTGATACCTCTCTCTTGGTGGTGGGTTACACAGCGCGCAGCAAT 1413	
Db	1298	CTTCTGGAAGCTCTTACAGCTCTCATTTGGTGGAAACAGGCTATGAGCGCGGACACAAAC 1357	
QY	1414	TCAGGAGGTCAATGGCTGCAGACCGTGACCGTTGGTTCGGATGGAATATGCTCTGTTT 1473	
Db	1358	TGACTGAGATGTTCTCTGCACACCGTGACCGTGGGCTCAGACAAAAGGTCCCGCTTT 1417	
QY	1474	CTATGSCAGTGGGCTACCTAGGTTATCTATCCGACTCAGAAAGTTGGCAGGTAGCAAGA 1533	
Db	1418	CCATGCTAGTGGTTGCCCGGGTGTTCCTCCAAACGCTGGGCTGAACGGAAGTACTG 1477	
QY	1534	TCTGTAGTAG 1543	
Db	1478	TTTGTACTTG 1487	
RESULT 9			
ACC43088			
ID	ACC43088	standard; DNA; 1485 BP.	
XX	ACC43088;		
XX	17-JUN-2003	(first entry)	
DE	Nucleotide sequence of an alpha-amylase of <i>Aspergillus niger</i> .		
KW	Alpha-amylase; enzyme; endohydrolysis; 1,4-alpha-glucosidic linkage;		
OS	oligosaccharide; polysaccharide; baking; ss.		
OS	<i>Aspergillus niger</i> .		
PN	WO2003016535-A2.		
PD	27-FEB-2003.		
PF	02-AUG-2002; 2002WO-NL000522.		
PR	16-AUG-2001; 2001EP-00000379.		
PR	16-AUG-2001; 2001EP-00000380.		
PR	16-AUG-2001; 2001EP-00000381.		
PR	16-AUG-2001; 2001EP-00000382.		
PR	16-AUG-2001; 2001EP-00000383.		
PR	16-AUG-2001; 2001EP-00000384.		
XX	(STAM) DSM NV.		
PI	Maier D, Stock A, Wagner C, Folkers U, Albermann K, Hopper S;		
XX	WPI; 2003-312758/30.		
XX	New polynucleotide from a filamentous fungus, preferably <i>Aspergillus</i>		
PT	<i>niger</i> , useful in a baking process.		
XX	Claim 1; Page 54-55; 81pp; English.		
XX	The present sequence encodes an alpha-amylase enzyme of the filamentous		
CC	fungus <i>Aspergillus niger</i> . Alpha-amylases catalyse the endohydrolysis of		

CC 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.
CC The alpha-amylase polynucleotides and polypeptides of the invention are
CC useful in a baking process
XX
SQ Sequence 1485 BP; 389 A; 430 C; 323 G; 343 T; 0 U; 0 Other;
Query Match 43.8%; Score 759.8; DB 8; Length 1485;
Best Local Similarity 70.0%; Pred. No. 4.9e-165;
Matches 1038; Conservative 0; Mismatches 442; Indels 3; Gaps 1;
QY 68 GTCCTATTTCGTAGGCGCTTCAGGTGCGGAGACCTGCTTTGGCTGCAACGCCCTGCGGA 127
DB 3 GACAAATCTTTCTGTTCTGCGCAATTTCTGCTGCTACAGCTCTGGCAGCCAGCTGCAGA 62
QY 128 CTGGGATCGCAATCATTATTTCTCTCACGGATCGATTTCGAAGACGAGTGGGTC 187
DB 63 ATGGGCTCCCAAGTCGATATATTTCTGCTCACGATCGCTTTGGCGNACGGATAATTC 122
QY 188 GACGACTGCGACTTGTATPACTGCGGATCAGAAATATCTGTGTGGAACATGCGAGGCAT 247
DB 123 TACCACCTGCTTCTGTGACTTGAGCGCTCGGCAATATTGCGGTGGATCTGCGAGGGCAT 182
QY 248 CATCGAATGTGGACTATATCATCAGGGAATGGCTTCACAGGCATCTGGATCACCCCGT 307
DB 183 CATCAATCAGCTGGACTATATTTCAAGGAATGGCTTTTACAGCGATCTGGATCACACCCGT 242
QY 308 TACAGCCAGCTGCCCCAGACACCGCATATGAGATGCTTACCATGGCTACTGCGCAGCA 367
DB 243 AACTGACAGATCCCCCAAGATCTGTTACGGACAGGATATCACGGNATCTGGCAGCA 302
QY 368 GGATATATATCTCTGAAAGAAATACCGGCACTGCGATGATGATTTGAAGGCGCTCTCTTC 427
DB 303 GGAGCGTTATGCCCTGAACTCCCATTTATGTTAGCGCAGACGATCTCAAAGCTCTGGCTTC 362
QY 428 GGCCTTCATAGAGGGGATGTATCTTATGTGCGATGTGTTGCTTAACCATATGGGCTA 487
DB 363 AGCTCTTCACTCACGGGCGATGATCTCATGTGGACGCTTGTGGCAATCATATGGGCA 422
QY 488 TGAATGAGCGGTGAGCTCAGTTCGATTTACAGTGTGTTAAACCGTTTCAGTTCCCAAGACTA 547
DB 423 CAATGTGACGGGAGCTCTGTGGACTCTGTGATTTACAATTAACAGACAAAGCTTGAAGACTGCTG 482
QY 548 CTTCCACCGCTCTGTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGGATTTGCTG 607
DB 483 CTTTCACACCTCTGTTGGATCTCTGATTTACAATTAACAGACAAAGCTTGAAGACTGCTG 542
QY 608 GCTAGAGATATACATGTCCTCTGCTGATCTCGATACCAACAGGATGTGGTCAAGAA 667
DB 543 GCTAGGCGATAACACCGTTGCGCTTGGCGATCTTGATATACAGTACGAGGTTGAAGAA 602
QY 668 TGAATGGTACGACTGGGTGGATCATTTGTTATCGAATCTCCATTCAGCGGCTCGGTAT 727
DB 603 TATGTGTTATGACTGGGTGAGTCTCTGCTCTTAACACTCTCGTGGCGGCTCGGCT 662
QY 728 CGACACAGTAAACACGTCAGAAAGACTTCTGGCCCGGTTACAAAGCCGCGAGGCT 787
DB 663 AGACACAGTCAAGACGTACAGAAAGACTTCTGGCCCGGCTACAAATGCTTCAGGCT 722
QY 788 GTACTGTATCGGGAGGTGCTGACGGTGTATCGGCTTACATCTGCTTACCAAGACGT 847
DB 723 GTACTGTATGGAGAAGTCTCGATGGGACGCTCATACAGCTGTCTTATCAGGAAGA 782
QY 848 CATGAGCGGCTGAACTATCCATTTACTATCCACTCTCAAGCGCTTCAAGTCAAC 907
DB 783 CTTGACCGAGTCTTAAATACCCCATGTACTATCCACTCTCGGCTCTTCAATCCAC 842
QY 908 CTTCCGCGACATGGACGCTCTTACAAATGATCAACACCGTCAATTCGACTGTCCAGA 967
DB 843 CAACGCGACTATCAGGACCTCTATTAACATGATCAACACCGTGAATCCACCTGCGAGA 902
QY 968 CTCACACATCTCGGGCACAATTCGTGAGAAACCAACGCAACCCACGGTTCGC---TTCTTA 1024
DB 903 TTCTACGCTTCTAGGACCTTGTGCGAAACCAACGATTAACCCACGCTTTGCCAAAGCTA 962

QY 1025 CACCAACACATAGCCCTCGCCAGAAAGCTGCGAGCATTCATCTCTCAACGACGGAAT 1084
DB 963 CACAAGCGACATGTCCCTAGCCAAAATGCCGCAACATTTCACTATCTCTGGCTGACGGCAT 1022
QY 1085 CCCCATCATCTAGCCGCGCAAGAACAGACACTACGCCGCGGAAACGACCCCGCAACCG 1144
DB 1023 TCCCATCATATACGCGCTCAGGAACAGCACTATATAGCGCGGTAATGACCCCTCAACACG 1082
QY 1145 CGAGCAACCTGGCTCTCGGCTACCCGACGACGAGCTGTACAAGTTAATTTGCCCTC 1204
DB 1083 CGAAGCGACCTGGCTTTTTCAGGCTACAAGACCACGAGGCTCTACCCATATCGCCGC 1142
QY 1205 CGCAACCGCAATCCGGAATATGCCATTAGCAAAAGATACAGGATTCGTGACCTACAAGAA 1264
DB 1143 ATCGAACAAGATTCCGACCCAGCTATTAACAGGATACCGGATATCTCACCTACAAAA 1202
QY 1265 CTGGCCCATCTTACAAGACGACAAACGATCGCCATCGCAAGGGCAGATGGGTGCGA 1324
DB 1203 CTACCCCATCTACCAAGACACCTCGACCTTTGCCATGGCAAGGCTACAATGGCACCCA 1262
QY 1325 GATCGTGAATCTTGTCCAAACAGGCTGCTCGGGTATTCGTATACCTCTCTCTTGAG 1384
DB 1263 AACTATCAAGCTCTTTTCTAACCTTGGCGCTCTCGGGCTCTCATACACTCTCCCTCCC 1322
QY 1385 TGGTGGCGGTTACACAGCGCGCCAGCAATTTGACGAGGCTCATTTGGCTGTCACGACCGTGAC 1444
DB 1323 AGGAACAGGCTACACAGCGCGCCAAAGATTTACTGAATCTATCTGACCAAGATCTAAC 1382
QY 1445 GGTGGTTCCGATGAAATGTGCTGTTCTTATGCGAGGTGGGCTACCTAGGATTTGTA 1504
DB 1383 AGTCAACTCAATGGCTCGGTGCCAGTACCCATGAAGAGCGGTTTACCGCGGATCTCTTA 1442
QY 1505 TCCGACTCAGAGTTGGCAGGTAGCAAGATCTGTAGTAGTCG 1547
DB 1443 TCCTGCAGATAAGTTGGTTAATGGAAGCTCATTTTTCAGTTAG 1485
RESULT 10
ABQ80348
ID ABQ80348 standard; cDNA; 1893 BP.
XX
AC ABQ80348;
XX
DT 11-AUG-2003 (first entry)
XX
DE A. fumigatus AfAAL2 cDNA.
XX
KW Gene; enzyme; tannase; cellulase; glucose oxidase; glucoamylase; phytase;
KW beta-galactosidase; invertase; lipase; alpha-amylase; laccase;
KW polygalacturonase; xylanase; gallate ester linkage; detergent; cellulose;
KW glucose; oxygen; myo-inositol phosphate; lactose; tea leaf; sucrose;
KW glyceride; starch; maltodextrin; oxidated phenolic compound;
KW polygalacturonic acid chain; xylan; xylo-oligomer; food; feed; beverage;
KW textile; tea liquor; cleaning ability; ss.
XX
OS Aspergillus fumigatus.
XX
FH Key Location/Qualifiers
CDS 1..1893
FT /*tag= a
FT /product= "Alpha amylase 2"
XX
FN WO2003012071-A2.
XX
PD 13-FEB-2003.
XX
XX 05-AUG-2002; 2002WO-US024842.
XX
XX 03-AUG-2001; 2001US-0309870P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX

PI Jiang B, Storms R, Roemer T, Bussey H;
XX WPI; 2003-332729/31.
DR P-PSDB; ABB80178.
XX
PT Novel isolated Aspergillus fumigatus polypeptide, useful in various
PT industries such as those involved in the making of food and feed,
PT beverages, textiles and detergents.
XX
PS Claim 2; Page 136-39; 169pp; English.
XX
XX The sequences given in AB080319-66 encode enzymatic proteins derived from
CC A. fumigatus. The resulting proteins display the catalytic activity of an
CC enzyme such as tannase, cellulase, glucose oxidase, glucoamylase, laccase,
CC phytase, beta-galactosidase, invertase, lipase, alpha-amylase, laccase,
CC polygalacturonase or xylanase. Compositions comprising the tannase are
CC useful for modulating the amount of compounds that comprise a gallate
CC ester linkage in a composition. Compositions comprising cellulase are
CC useful for modulating the amount of cellulose in a composition.
CC Compositions comprising glucose oxidase are useful for modulating the
CC amount of glucose or oxygen in a composition. Compositions comprising
CC phytase are useful for modulating the amount of myo-inositol phosphates
CC in a composition. Compositions comprising beta-galactosidase are useful
CC for modulating the amount of lactose in a composition. Compositions
CC comprising sucrose or invertase are useful for modulating the amount of
CC sucrose in a composition. Compositions comprising lipase are useful for
CC modulating the amount of glyceride in a composition. Compositions
CC comprising alpha-amylases are useful for modulating the amount of
CC starches or maltodextrins in a composition. Compositions comprising
CC laccase are useful for modulating the amount of oxidized phenolic
CC compounds in a composition. Compositions comprising polygalacturonases
CC are useful for modulating the amount of high or low molecular weight
CC polygalacturonic acid chains in a composition. Compositions comprising
CC xylanases are useful for modulating the amount of xylan or xylo-oligomers
CC in a composition. The A. fumigatus proteins and corresponding DNA's are
CC useful in various industries such as those involved in the making of food
CC and feed, beverages, textiles and detergents. The DNA's are useful to
CC express recombinant enzymes for characterization, modification or
CC industrial uses, to compare with the nucleotide sequence of A. fumigatus
CC to identify duplicated genes of paralogs having the same or similar
CC biochemical activity and/or function, to compare with nucleic acid
CC sequence of other related or distant fungal organisms to identify
CC potential orthologous enzyme genes, for selecting and making oligomers
CC for attachment to a nucleic acid array for examination of expression
CC patterns, and to raise anti-protein antibodies. The polypeptide having
CC tannase activity increases the yield of tea liquor from tea leaves,
CC improves the colour, flavour and health benefits of tea products,
CC particularly an instant tea product. The polypeptide having cellulase
CC activity enhances cleaning ability of detergent compositions
XX
SQ Sequence 1893 BP; 425 A; 568 C; 475 G; 425 T; 0 U; 0 Other;
Query Match 37.4%; Score 648.2; DB 10; Length 1893;
Best Local Similarity 65.4%; Pred. No. 2.8e-139;
Matches 966; Conservative 0; Mismatches 508; Indels 3; Gaps 1;
73 TATTCTGTAGCGCCTTCAGGTCGCGCACCTGCTTTGGCTGCAACGCGCTCGGACTGGC 132
29 TGTCCCTGTGCTGCTCCCTCTCGGACAGGCTGCCATGCTCTGACCCGAGCGGATGGC 88
133 GATCGGAATCCATTATTTCCTCTCAGCGATCGATTGTCGAAAGGAGCGGATGGGTCGACGA 192
89 GCAGTCAATCGATCTATTTCTCTCTGACCGATCGGTTGCGGCCGAGAAGACAAATCCACGA 148
193 CTGGGACTGTATCTACTGCGGATCAGAAATCTGTGTGGAATCTGCGAGGCGATCATCG 252
149 CTGCTGCTCGGATGTCACCAACGACTGTATTGCGGCGGGAGCTGGCAGGGGATCATCA 208
253 ACAAGTTGGACTATATCCAGGGAATGGCTTTCACAGCCATCTCGATCAACCCCGTTACAG 312
209 ATCATCTGACTACTATCAAGGATGGGATTTACTGCCATATGATCAACCCCGTAACTG 268
313 CCAGCTGCCCCAGACCAACCGCATATGGAGATGCCTACCATGGCTACTGGCAGCAGGATA 372

269 AGCAGTTCTATGAGAAACACCGCGCATGGTACTTCTGTAACATGATACTGGCAGCAGAATA 328
373 TATACTCTCTGAAACGAAACACTACGGCACTGACAGATGACTTGAAGGCGCTCTCTTCGGCCC 432
329 TCACAGGAGTCAATGCCATTATGGAACGGCACAAGATCTTAGAGATCTGCGCCACGCTC 388
433 TTATGAGAGGGGATGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 492
389 TGCACGCGGCTGCGCATGTACTTGTATGATGATGATGATGATGATGATGATGATGATGATG 448
493 GACGCGGTAGCTCAGTCGATTAACAGTGTGTTAAACCGTTTCAAGTTCCTCCCAAGACTACTTC 552
449 GACGCGGAAACTCGGCAACTACCGGTGCTTCTACTCGTTTGAATTCGCTACCTATTATTC 508
553 ACCCGTTCTGTTTTCATTTCAAACTATTAAGATCAGACTCAGGTGAGGATTTGCTGGCTAG 612
509 ACCATACTGTCTCATCACCAGCTACAACAACCAACAGCTGTGGAGGACTGCTGGCTGG 568
613 GAGATAACTGTCTCTCTTCTGCTGATCTCGATACCAACCAAGGATGGTCAAGATGAAT 672
569 GAGATACTACTGTCTGCTACCGATCTCGACACGACGACGCGGAGTGGCAAGCATCT 628
673 GGTACACTGGGTGGGATCATTTGGTATCGAATCTACTTCCATTTGACGGCTCCGTATCGACA 732
629 GGTATGATTGGGTGAAGGATTTGGTTCGCAACTACTTCCATTCGACGGCTCCGTATCGACA 688
733 CAGTAAACACGCTCCAGAAAGGACTTCTGCGCGGGTACAACAAGCGCGAGGGGTGTACT 792
689 CGGTGAAGCATGTGGAAGAAGACTTCTGCGCGGCTACATGATGAGCTGTGGGTCTACT 748
793 GTATCGCGGAGGTGCTGACGCGGTATCTCGGCTTACATTTGTCTCTACCAAGAGCTATGG 852
749 GTGTCGGTGAAGTCTTTTTCGGGTGATCCACAATATACCTGTCCATACCAAGATTTACCTGG 808
853 ACGGCTGCTGACTACTTCCATTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 912
809 ATGGTGTACTTAACATCCCATATATCTCAACTTCTCTACGGGTTCCTCAATTCAGCAGCG 868
913 GCAGCATGACGACCTCTCAACATGATCAACACCGCTCAAAATCCGACTGTCCAGACTCAA 972
869 GCAGCATGACGATCTGTACACATGATCAGCTCCGTTGGTGTGACTGTGGGATCCCA 928
973 CACTCTGCGGCACTTCTGCGAAGAACGACGACCAACCGGTTGCTTTTACACCAAG 1032
929 CTTTGTCTGCGCACTTTATCGAAGAACATGATTAACCCCGATTTGCTCTCTATACGAGCG 988
1033 ACATAGCCCTCGCAAGACGTCGCGAGCATTCATCTCTCAACGAGGATCCCATCA 1092
989 ACTATTGCGAAGCCAAAGACGTCATCTCTCTCATGTTTCTCTCCGAGGATCCCATTTG 1048
1093 TCTACGCGCGCAAGAACGACGACTACGCGCGGGAACGACCGCGGAACCGGAGCAA 1152
1049 TCTACGCGCGCAGGAGGAGCATCTACGCGCGGCTGCTGACCTGCAACCGGAGGCTG 1108
1153 CTGCGCTCTGCGGCTACCGGACCGACGAGAGCTGTACAGTTTAATTTGCTTCGCGAAGC 1212
1109 TCTGCGCTGTGATACTCGACCGGCTACGCTGTACAGTGGATTGCTCTTACCAACA 1168
1213 CAATCGGAATCTATGCCATTAGCAAGATACAGGATTCGTGACCTACAGAACTGGGCCA 1272
1169 AGATTGCGAAGCTAGCGATTTCCAAAGACTAGCCTACATACATCAAGAAACAACCGCT 1228
1273 TCTACAAACGACACACAACGATCGCCATCGCAAGGGGACA --- GATGGTTCGAGATCG 1329
1229 TCTATGATGATTCGAATCTCTCGCTATGCGCAAGGGGCTCAGTCGCTGGCTCTCAAGTCA 1288
1330 TGACTATCTTGTCCAAAGAGGTGCTTCGGGTGATTCGTATACCTCTCTCTTGGTGGTG 1389
1289 TTACCGTCTCAGTAACAAGGGATCTCGGGCAGTTCTTACACCTCTCTCTCTCAGCGGCA 1348
1390 CGGTTTACACAGCGGCGCAGCAATTTGAACGAGGTCTATTGGCTGACAGCCGCTGACGCGTTG 1449

Db 1349 CGGGCTACTCCCGCGCGCCACCCCTTGTTCGAGATGTATACATGCACTACTCTCAACCGTGG 1408
Qy 1450 GTTTCGATGGAATGTGCTGTTCCTATGCGCAGGTGGCTACTAGGGTATTGTATCCGA 1509
Db 1409 ACTCGAGCGGAATCTCGCGGTGCCNATGGTATCCGGCTTGCCAGAGTTTCGTGCGCT 1468
Qy 1510 CTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTC 1546
Db 1469 CGTCATGGGTCACTGGAGTGGCCCTCTCGCGGCACTC 1505

RESULT 11
ID AAQ88712
XX AAQ88712 standard; cDNA; 1956 BP.
AC AAQ88712;
XX 27-AUG-2003 (revised)
DT 05-JAN-1996 (first entry)
XX Aspergillus kawachii (IFO 4308) alpha-amylase.
XX
XX Aspergillus kawachii; IFO 4308; alpha-amylase; alcohol;
KW starch saccharification; transformed Saccharomyces cerevisiae;
KW recombinant production; acid resistant; heat stable; ss.
XX
XX Aspergillus kawachii.
XX
XX JP07099979-A.
PN
XX
XX 18-APR-1995.
XX
XX 30-SEP-1993; 93JP-00265447.
XX
XX 30-SEP-1993; 93JP-00265447.
PR
XX
XX (KOKU-) KOKUZEI CHO CHOHAN.
PA (JOZO-) JOZO SHIGEN KENKYUSHO KK.
XX
XX WPI; 1995-182071/24.
DR
XX
XX Aspergillus kawachii alpha-amylase gene and transformant - for
PT recombinant prodn. of acid-resistant, thermostable enzyme useful for
PT direct saccharification of starch to produce alcoholic beverages.
PS
PS Claim 1; Page 5; 7pp; Japanese.
XX
XX AAQ88712 encodes the A. kawachii (IFO 4308) alpha-amylase, it was used in
CC the construction of an expression vector for the transformation of a S.
CC cerevisiae host. The transformed S. cerevisiae produced recombinant alpha
CC -amylase, which was pH (pH 2-7) resistant and heat stable (80% at 60 deg.
CC C for 60 mins.). The alpha-amylase allows the direct saccharification of
CC starch for the prodn. of alcohol, and alcoholic beverages. (Updated on 27
CC -AUG-2003 to correct OS field.)
XX
XX Sequence 1956 BP; 457 A; 569 C; 485 G; 445 T; 0 U; 0 Other; 1

Query Match 32.9%; Score 570; DB 2; Length 1956;
Best Local Similarity 64.1%; Pred. No. 3e-121;
Matches 910; Conservative 0; Mismatches 450; Indels 60; Gaps 1;

Qy 121 CTGCGGACTGGCGCATCGCAATCCATTTATTTCTCTCAGCGATCGATTGTGCAAGGACGG 180
Db 71 CTGCAAGTGGCGCATCTCAATCATCTACTTCTTTTGCGGATCGGTTGCGTAGGACGG 130
Qy 181 ATGGGTTCGACGACTGGCAGTTGTAATCTGCGGATCAGAAATCTGTGGTGAACATGGC 240
Db 131 ACAATTCGACTACAGCTACGTGCAATACGGGTGACCAAAATCTACTGTGTGGAAGTTGGC 190
Qy 241 AGGGCATTCGACAGTTGGACTATATCCAGGGAATGGCTTCAGGCCATCTGGATCA 300
Db 191 AAGGAATTATCAACCATCTGGACTATATCCAGGGCATGGGATTCAGACTATCTGGATCT 250

Qy 301 CCCCGTTTACAGCCAGCTGCCCCCAGACCAACCGCATATGGAGATCGCTACCATGGCTACT 360
Db 251 GCGCTATCACTCAGCAGCTAGCCCCAGGATACTTCCGATGGTGAAGCCTACCATGGATACT 310
Qy 361 GGCAGCAGGATATATCTCTCTGACGAAACTAGCGCACTGCAGATGACTTGAAGGCGC 420
Db 311 GGCAGCAGGATATATCTATGTGAATCTCAACTTCGGCAGCGCAGATGATCTGAAGTCCC 370
Qy 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTTATGTCGATGTGTTGTCATAACCAT 480
Db 371 TCTCGATGCTCTTCACGCCCGCGAATGTACTCATGTCGACGTCGTCCCTTACCACA 430
Qy 481 TGGGCTATGATGGAGCGGTAGCTCAGTCGATTAAGTGTGTTTAAACCGTTTCAGTTCCC 540
Db 431 TGGGCTACGCGAGGTAAACGCAACGATGTGGATTACAGGCTCTTCGACCCCATC----- 483
Qy 541 AAGACTACTTCCACCCGTTCTGTTTTCATTCAAACTATGAAGATCAGACTCAGGTTGAGG 600
Db 484 -----GTCCAG 490
Qy 601 ATTGCTGCTAGGAGATAACACTGTCTCTTGCCTGATCTCGATACCACCAAGGATGTGG 660
Db 491 ACTGTTGGAGGCTGACACCATCTGTCTCTGCCAGATCTGAACACCAAGGAAACCGCG 550
Qy 661 TCAAGAATGAATGGTACGACTGGGTGGATCAATGGTATCGAACTACTCCATTTGACGGCC 720
Db 551 TGAGAAACCATTTGGTACGATTGGGTAGCGGACTGGTATCCAACTACTCAGTCGACGGCC 610
Qy 721 TCCGTATCGACACAGTAAACACGTCGAGAGGACTTCTGGCCCGGTACAAACGCG 780
Db 611 TCCTGATCGACAGTGTGAAGAAGTCGAACCCGAGTTCTTCCCGGCTACGAAGAAGCAG 670
Qy 781 CAGCGGTGTACTGTATCGCGAGGTCTCGACGGGTGATCCGGCCTACACTTGTCCCTACC 840
Db 671 CAGGAGTCTACTGCGTGGTGAAGTCGAACGCGCAACCTGCTCTCGACTGCCATACC 730
Qy 841 AGAAGCTCATGACCGCGTACTGAATCTATCCATTATCTATCCACTCTCTCAACGCTTCA 900
Db 731 AAAAATATCTAGATGGTGTCTCAACTATCCCATCTACTGGCAACTCTCTACGCGCTTG 790
Qy 901 AGTCAACCTCCGGCAGCATGGACGCTCTACAAATGATCAACACCGTCAAAATCGACT 960
Db 791 AATCTCCAGCGCGCAGCATCAGCAACCTCTCAAAATGATCAAAATCCGTCGCGAGCAT 850
Qy 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGCGAACCACGCAACCCACCGTTCGCTT 1020
Db 851 GCTCGATCCGACCTCTCTGGCACTTTATCGAANAACACGCAACCCCGCTTGCCT 910
Qy 1021 TTACACCAACGATATAGCCCTTCGCCAAGAACTGCGCAGCATTCATCTCTCAACGAG 1080
Db 911 CCTACACATCCGACTACTCCCAAGCCMAAAGTTCTCAGCTACATCTTCTCTCCGAGC 970
Qy 1081 GAATCCCATCATCTACGCCGCGCAAGACAGCATCTACGCCGCGGAAACGACCCCGCA 1140
Db 971 GCATCCCATCTGCTACGCCGCGCGGAAGACAGCATCTCTCCGCGGAGAGTGCCT 1030
Qy 1141 ACCGGAAGCAACCTGGCTCTCGGCTACCCGACGACAGCGAGCTGTACAAAGTTAAATTG 1200
Db 1031 ACAACGGAGTACTTGGCTATCAGGTACGACACCTCCGCGAGCTCTACACTGGATAG 1090
Qy 1201 CCTCCGCAACGCAATCCGAACTATGCGATTAGCAAGATACAGGATTCTGACCTCA 1260
Db 1091 CCACCACAAACCGATCCGAACTAGCTATCTCAGCAGACTCGGACTACATTACTTACA 1150
Qy 1261 AGAATGGCCCATCTACAAAGACGACAAACGATGCCATGGCAGGCGCAGATGGGT 1320
Db 1151 AGAAGCAGCCCAATCTACACAGACGACCAACCATCGCATGGCAAGGCACTCCGGCT 1210
Qy 1321 CCAGATCTGACTATCTTGTCCAAAGGGTGTCTCGGGTGTATTCGTATACCTCTCCT 1380
Db 1211 CCCAATCATCACCGTCTCTCCAAACAGGCTCTCCGGAAGCAGCTACACCTCACCC 1270
Qy 1381 TGAGTGTGCGGGTTACACAGCCGCGCAGCAATTAAGCGAGGTCTATTGGCTGACAGCG 1440

Db 1271 TCAGCGAGCGGCTACACCTCGGCACGAGCTCATCGAAGGTATACCTGCGCTCCG 1330
QY 1441 TGACGGTGTGCTCGGATGGAATGTGCTCTTCTATGGCAGGTGGCTACCTAGGGTAT 1500
Db 1331 TGACGGTGGACTCGAAGCGGGATATCCCTGTGCGGATGGCTTCGGGATTACCTAGAGTTC 1390
QY 1501 TGATTCGACTGAGAGTTCGGCAGGTAGCAAGATCTGTAG 1540
Db 1391 TCCTCCCTGCTCGGTGGTTGATAGTTCCTTCGCTTTGTGG 1430

RESULT 12

AAFI1252/c

ID AAF11252 standard; cDNA; 1389 BP.

XX AC

AAF11252;

XX AC

DT 13-MAR-2001 (first entry)

XX DE

Aspergillus niger EST SEQ ID NO:3775.

XX DE

Multiple gene expression; filamentous fungal cell; EST;

XX KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;

XX KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;

XX KW culture condition; environmental stress; spore morphogenesis;

XX KW metabolic pathway engineering; catabolic pathway engineering; ss.

XX OS

Aspergillus niger.

XX XX

XX PN WO200056762-A2.

XX PN

XX PD 28-SEP-2000.

XX XX

XX PF 22-MAR-2000; 2000WO-US007781.

XX XX

XX PR 22-MAR-1999; 99US-00273623.

XX XX

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PA

XX PA (NOVO) NOVO NORDISK AS.

XX XX

XX PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

XX XX

XX DR WPI; 2000-594572/56.

XX XX

XX PT Monitoring differential expression of genes in filamentous fungal cells

XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a

XX PT substrate of expressed sequence tags.

XX XX

XX PS Claim 87; Page 1694-1695; 3161pp; English.

XX XX

XX CC The present invention describes a method for monitoring differential

XX CC expression of genes in a first filamentous fungal (FF) cell relative to

XX CC expression of the same genes in one or more second filamentous fungal

XX CC cells. The method uses fluorescence-labeled nucleic acids isolated from

XX CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs

XX CC are used in the methods for monitoring differential expression of genes

XX CC in a first filamentous fungal (FF) cell relative to expression of the

XX CC same genes in one or more second filamentous fungal cells. Monitoring the

XX CC global expression of genes from FF cells allows the production potential

XX CC of the microorganisms to be improved. New genes may be discovered,

XX CC possible functions of unknown open reading frames can be identified and

XX CC gene copy number variation and stability can be monitored. The expression

XX CC of genes can be used to study how FF cells adapt to changes in culture

XX CC conditions, environmental stress, spore morphogenesis, recombination,

XX CC metabolic or catabolic pathway engineering. Using ESTs provides several

XX CC advantages over genomic or random cDNA clones including elimination of

XX CC redundancy as one spot on an array equals one gene or open reading frame,

XX CC and organisation of the microarrays based on function of the gene

XX CC products to facilitate analysis of the results. AAF07478 to AAF1247

XX CC represents ESTs from Fusarium venenatum; AAF1248 to AAF1853 represents

XX CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from

XX CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from

XX CC

CC Trichoderma reesei, which are all specifically claimed in the present
XX invention

SQ Sequence 1389 BP; 354 A; 384 C; 325 G; 326 T; 0 U; 0 Other;

Query Match 25.3%; Score 439; DB 3; Length 1389;

Best Local Similarity 98.9%; Pred. No. 4.3e-91;

Matches 442; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1174 CGCAGCGAGCTGTACAAAGTTAAATTCCTCCGCGAACCAATCCGGAACTATGCCATTA 1233

Db 1364 CCGCGCTAGTCTGTACAAAGTTAAATTCCTCCGCGAACCAATCCGGAACTATGCCATTA 1305

QY 1234 GCAAAGATACAGGATTTCGTGACCTACAAGAACTGGCCCATCTTACAAGACGACACACGA 1293

Db 1304 GCAAAGATACAGGATTTCGTGACCTACAAGAACTGGCCCATCTTACAAGACGACACACGA 1245

QY 1294 TCGCCATGCGCAAGGCGACAGATGGGTGCGAGATCGTGACTATCTTTGTCCCAAGGGTG 1353

Db 1244 TCGCCATGCGCAAGGCGACAGATGGGTGCGAGATCGTGACTATCTTTGTCCCAAGGGTG 1185

QY 1354 CTTCCGGTGATTTCGTATACCCCTCTCCTTGAGTGGTGGGTTACACAGCGCGCAGCAAT 1413

Db 1184 CTTCCGGTGATTTCGTATACCCCTCTCCTTGAGTGGTGGGTTACACAGCGCGCAGCAAT 1125

QY 1414 TGACGAGGTCAATTCGTGCTGCAGCCGTCACCGTTCGGTTCGGATGGAATGTGCTGTTTC 1473

Db 1124 TGACGAGGTCAATTCGTGCTGCAGCCGTCACCGTTCGGTTCGGATGGAATGTGCTGTTTC 1065

QY 1474 CTATGCGAGTGGGCTACCTAGGTTATTTATCCGACTGAGAGTTGGCAGGTAGCAAGA 1533

Db 1064 CTATGCGAGTGGGCTACCTAGGTTATTTATCCGACTGAGAGTTGGCAGGTAGCAAGA 1005

QY 1534 TCTGTAGTAGCTCGTGAAGGTTGGAGATATATGATGGTACTGCTATTCAATCTGGCATT 1593

Db 1004 TCTGTAGTAGCTCGTGAAGGTTGGAGATATATGATGGTACTGCTATTCAATCTGGCATT 945

QY 1594 GGACAGTGAGTTTGAGTTTGATGTACA 1620

Db 944 GGACAGTGAGTTTGAGTTTGATGTACA 918

RESULT 13

AAQ25079

ID AAQ25079 standard; DNA; 1404 BP.

XX XX

AC AAQ25079;

XX XX

DT 25-MAR-2003 (revised)

DT 17-NOV-1992 (first entry)

XX XX

DE Alpha-amylase variant encoding leucine at position 84.

XX XX

XX KW Mutant; maltose; malto-oligosaccharides; Saccharomycopsis; fibuligera;

XX KW polymerisation; DP; transglycosifier; ss.

XX XX

OS Saccharomycopsis fibuligera.

XX XX

FH Key Location/Qualifiers

FT mutation

FT 329

FT /*tag= a

FT /note= "mutated to thymine"

XX XX

PN JP04108386-A.

XX XX

PD 09-APR-1992.

XX XX

PF 28-AUG-1990; 90JP-00226112.

XX XX

PR 28-AUG-1990; 90JP-00226112.

XX XX

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

XX XX

DR WPI; 1987-173694/25.
DR P-PSDB; AAP70571.
XX Amylase prodn. - comprises culturing microorganism transformed with
PT vector deoxyribonucleic acid, accumulating and collecting amylase.
XX PS
XX Disclosure; Fig 1; 14pp; Japanese.
XX
CC The plasmid may be used to transform an E.coli expression system for the
CC stable production of amylase, useful in ethanol fermentation. See also
CC AAN70917. (Updated on 24-OCT-2003 to standardise OS field)
XX
XX Sequence 4214 BP; 1249 A; 784 C; 860 G; 1321 T; 0 U; 0 Other;
SQ
Query Match 21.8%; Score 377.4; DB 1; Length 4214;
Best Local Similarity 54.8%; Pred. No. 9.1e-77;
Matches 778; Conservative 0; Mismatches 626; Indels 15; Gaps 1;
55 TGGTCCGGTGGTGTCTCTATTCTGTACGGCTTACGGTCGGCACCTGCTTTGGCTG 114
1553 TGGTGGCTCATTTGGCTGGCCCTTGTATTGCTCAACAGTACTCTATTCAAAGAGAAA 1612
115 CAACGGCTCGCGATCGCGATCGCAATCCATTTATTCTTCTCACGGATCGATTGGCAA 174
1613 CTAATGCTGATAAATGGAGATCACAGTCTATTATCAAAATGTCACGAGATTGCTA 1672
175 GGAAGGATGGGTGACGACATCGGACATTTGTAATCTGCGGATCAGAAATATCTGTGGTGA 234
1673 GAACCGATGGTGATACAAAGTGTCTCTGTAAACAGAGATAGACTTTTACTGTGGTGT 1732
235 CATGGGAGGCGATCATCGCAAGTGTGACTATATCCAGGGAATGGGCTTCACAGCCATCT 294
1733 CTTTCAAGGCATCATAAAGTGTGATTACATCAAGATATGGGCTTTTACTGCTATTT 1792
295 GGATCACCCCGTTACAGCCAGCTGCCCGACAGCCCGCATATGAGATGCCATCCATG 354
'1793 GGATTTCTCCAGTTGTGAAACATATCCCGATACAGCATATGGTATGCTTATCATG 1852
355 GCTACTGGCAGAGGATATATCTCTCTGAAAGAAATACAGGACATGCGAGATGACTTGA 414
1853 GTTACTGGATGAAGAACATATACAAATTAATGAAACCTTTGGTACTGCTGATGATTGA 1912
415 AGCGCTCTCTCGGCCCTTCATGAGGGGGATGATCTTATGTTGCTGATGTTGCTA 474
1913 AGTCTTTGGCACAAGAAATTCACGATGCTGATATGTTTAAATGGTGGATATCGTTACCA 1972
475 ACCATATGGCTATGATGGAGCGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTCA 534
1973 ACCATTACGGCAGTATGGCAGTGGAGATAGTATCGATTACTCAGAGTACACCCCGTTCA 2032
535 GTTCCCAAGACTACTCCACCCGTTCTGTTTCAATCAAACTATGAAGATCAGACTCAGG 594
2033 ACGACCAAAAGTACTTCCATACTACTCTTATTTTCAAACTATGATGACCAAGCTCAGG 2092
595 TTGAGATTTGCTGGTAGGAGATAACATCTCTCTTCTGCTGCTGATCTCGATACCAACAGG 654
2093 TTCAAAGTTGCTGGGAGGTGACTCTCTCAGTTGCAATTTACCAGATTTGAGAACGAGATA 2152
655 ATGTGTTCAAGATGAATGTACGACTGGGTGGATCATTTGGTATCGAACTACTCCATTG 714
2153 CGACGTGGCTCAGTTTCAATTTCTGGTTAAAGATTTTGGTGAATTTACTCAATG 2212
715 ACGGCTTCGATCGACACAGTAAACACGCTCCAGAAAGGACTTTCTGGCCGGGTACAACA 774
2213 ATGGTTTAAGAAATTGATGTGCTAAACATGTGGACCAAGGCTTTTCCCGGATTTTGTGA 2272
775 AAGCGGAGCGGTGACTGTATCGGCGAGGTGCTCGACGGTGTATCGGCGCTACACTTGTG 834
2273 GTGCATCTGGAGTTTACTCAGTAGGGGAAGTTTTTCCAAGGAGACCCAGCTTTATACATGCC 2332
835 CCTACCAAGACGTCATGGAGCGGTACTGAACATATCCATTTTACTATCCACTCTCAACG 894
2333 CATACCAAAATTACATCCAGGGGTAGTAAATTATTCATTTGTTACTACCCCAACCCAGAT 2392

QY 895 CTTCAAGTCNAACCTCCGGCAGCATGGACGACCTCTACACATGATCAACACCGTCAAT 954
DB 2393 TTTTAAAACTACTGATTCAGTTCAGTGGATTGACTCAATGATTTCAAGCGTTGCTT 2452
QY 955 CCAGCTGTCCAGACTCAACACTCCTGGGCACATTTCTGAGAACACACGACCAACCCCGT 1014
DB 2453 CCAGTTGTTGGATCCAACTTTTGTGACAACTTTGTAGAAATACAGATAATGAAGGT 2512
QY 1015 TCGTTCTTTACACAAACGACATAGCCCTCGCAGAAACGTCGACGATTCATCATCTCA 1074
DB 2513 TCGTTTCAATGACCGACGACCAAGTTTGAATTTCTAATGCTATTGTCATTTCTCTTTGG 2572
QY 1075 ACAGCGAATCCCATCATCTACCGCGCCAGAACAGACACTACGCGCGGGAACGACC 1134
DB 2573 GTGATGGTATTCCTGTCTATTACTATGSAACAAAGCTTTGAGCGGAAAAGTGACC 2632
QY 1135 CCGCAACCGCAAGCAACCTGGCTCTCGGGCTACCCGACGAGCGAGCTGTACAAGT 1194
DB 2633 CAACACAGAGAGGCTTGTGGTTATCCGGCTACAAAGAGAGTACTATTACAGC 2692
QY 1195 TAATTCCTCCCGGAAACGCAATCCGGAATATGCCATTAGCAAGATACAGGATTCGTGA 1254
DB 2693 TCATTGCCAAAGCTAATGCTGCGAGAAACGCGCGCTTTATCAAGACTCAAGCTATGCCA 2752
QY 1255 CCTACAAGAACTGGCCCATCTACAAGACGACACACGATCGCCATCGCGCAAGGGCAG 1314
DB 2753 CTTGCGAGCTTTCTGTGATCTTTTCAAAATGACCATGTTATTGCAACAAAAGAGGCG 2812
QY 1315 ATGGGTGCGCAGATCGTGACTATCTTCTGCAAAAGGGTCTTCGGGTGATTTCGTATACC 1374
DB 2813 TGTGTTCTG-----TTTCAACAACTTGGTTCCAGGGTTCTTCTGATG 2857
QY 1375 TCTCTTGTAGTGGTGGGTTACACAGCGCGGACCAATTTGACGAGGTCAITGGCTGCA 1434
DB 2858 TGACTATTTCCAAACACAGGTTACAGTTCCGGTGAGGATTTGGTGAAGTTTTCACATGCA 2917
QY 1435 CGACCGTGACGGTGGTTCGGATGGAATGTGCTCTTC 1473
DB 2918 GTACTGTTAGCGCAGCTCTGACTTCAAGTTTCTATCC 2956
RESULT 15
AAQ77665
ID AAQ77665 standard; DNA; 1404 BP.
XX
XX AAQ77665;
XX AC
XX AC
DT 16-JUN-1995 (first entry)
XX
DE Variant alpha amylase deriv. from *Saccharomyopsis fibuligera*.
XX
XX alpha amylase; carbohydrate hydrolase; increased activity;
KW tyrosine residue; enzyme centre; mass production; oligosaccharide;
KW variant; cyclomalodextrin glucanotransferase; ds.
XX
XX *Saccharomyopsis fibuligera*.
XX
FH Key Location/Qualifiers
FT misc_difference 247..249
FT /tag= a
FT /note= "the wild type sequence TAY was mutated to TTC to
FT give a variant enzyme"
XX
PN JP06253836-A.
XX
PD 13-SEP-1994.
XX
XX 04-MAR-1993; 93JP-00069303.
XX
PR 04-MAR-1993; 93JP-00069303.
XX
PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2005, 18:46:17 ; Search time 323 Seconds
(without alignments)

8784.222 Million cell updates/sec

Title: US-10-820-200-1

Perfect score: 1734

Sequence: 1 tccatcaagctctcccttc.....aaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375.4	21.6	1404	1	US-08-204-656B-7
2	375.4	21.6	1404	1	US-08-470-702-4
3	375.4	21.6	1404	1	US-08-467-831-4
4	373.8	21.6	1404	1	US-08-204-656B-1
5	373.8	21.6	1404	1	US-08-204-656B-3
6	373.8	21.6	1404	1	US-08-204-656B-5
7	373.8	21.6	1404	1	US-08-470-702-1
8	373.8	21.6	1404	1	US-08-470-702-2
9	373.8	21.6	1404	1	US-08-470-702-3
10	373.8	21.6	1404	1	US-08-467-831-1
11	373.8	21.6	1404	1	US-08-467-831-2
12	373.8	21.6	1404	1	US-08-467-831-3
13	143	8.2	3311	4	US-09-367-891A-5
14	143	8.2	7432	4	US-09-367-891A-6
15	68.6	4.0	93	3	US-09-189-060B-63
16	68.4	3.9	919	3	US-09-189-060B-67
17	68.4	3.9	922	3	US-09-189-060B-65
18	68.4	3.9	922	3	US-09-189-060B-73
19	68.4	3.9	924	3	US-09-189-060B-69
20	63.4	3.7	912	3	US-09-189-060B-71
21	58	3.3	2582	2	US-08-816-105A-2
22	55.4	3.2	5163	3	US-08-700-651-1
23	55.4	3.2	5163	3	US-08-928-361B-4
24	55.4	3.2	5163	4	US-09-588-995A-4
25	55.4	3.2	5318	3	US-08-700-651-2
26	55.4	3.2	5318	3	US-08-928-361B-3
27	55.4	3.2	5318	4	US-09-588-995A-3

c 28 52.4 3.0 77536 4 US-09-410-551B-1 Sequence 1, Appli
c 29 52.4 3.0 77536 4 US-09-940-316B-1 Sequence 1, Appli
30 51.4 3.0 5511 3 US-08-928-361B-2 Sequence 2, Appli
31 51.4 3.0 5511 4 US-09-588-995A-2 Sequence 1, Appli
32 51.4 3.0 7334 3 US-08-928-361B-1 Sequence 1, Appli
33 51.4 3.0 7334 4 US-09-588-995A-1 Sequence 1, Appli
34 51.2 3.0 867 4 US-09-216-393B-340 Sequence 340, App
35 51.2 3.0 867 4 US-09-216-393B-342 Sequence 342, App
36 51.2 3.0 1397 4 US-09-216-393B-343 Sequence 343, App
c 37 51.2 3.0 1397 4 US-09-216-393B-345 Sequence 345, App
38 50.6 2.9 2160 3 US-09-386-607-1 Sequence 1, Appli
39 50.6 2.9 2160 4 US-09-645-707B-1 Sequence 1, Appli
40 49.2 2.8 1295 1 US-08-433-854-3 Sequence 3, Appli
41 49.2 2.8 1295 1 US-08-174-745A-3 Sequence 3, Appli
42 49.2 2.8 1295 2 US-08-195-947-3 Sequence 3, Appli
43 49.2 2.8 1295 2 US-08-433-885-3 Sequence 3, Appli
44 49.2 2.8 1295 2 US-08-433-908B-3 Sequence 3, Appli
45 49.2 2.8 1295 3 US-08-410-614-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-204-656B-7
; Sequence 7, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Derived from plasmid pSf1 (Agric. Biol. Chem.
FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 1..1404
; OTHER INFORMATION:
; /note= "Nucleotides 1-1404
; correspond to nucleotides 79-1482 of the Saccharomycopsis
; fibuligera '-amylase structural gene"
; OTHER INFORMATION:
US-08-204-656B-7

```

Query Match	21.6%	Score 375.4	DB 1	Length 1404
Best Local Similarity	55.6%	Prod. No. 2.9e-99		
Matches 752	Conservative 0	Mismatches 586	Indels 15	Gaps 1
QY	121	CTCGGACTGGCATCGCAATCCATTTATTTCTCTCAGGATCGATTTCGAAGACGG	180	
DB	11	CTGATAATGGAGATCACAGTCTATTTATCAATTTGCTACTGCAGATTTGCTAGAACGG	70	
QY	181	ATGGGTGCGACCTGGACCTTTTAATACTCGGGATCAGAAATACTGTGTGGAAACATGGC	240	
DB	71	ATGGTGATACAAGTGTCTCTGTAACACAGAAGATAGACTTTACTGTGTGGTGTCTTTTCC	130	
QY	241	AGGGCATCATCGCAAGTTGGACTATATCAGGGAATGGGCTTCACAGCCATCTGGATCA	300	
DB	131	AAGGCATCAAAAGAGTTGGATTACATAAAGATATGGGCTTTACTGCTATTTTGGATTT	190	
QY	301	CCCCGGTTACAGCCCCAGCTCCCCAGACCACCGCATATGGAGATGCCCTACATGGCTACT	360	
DB	191	CTCCAGTTGTTGAAAACATTTCCCGATAACACAGCATATGGTTATGCTTATCATGTGTAAT	250	
QY	361	GGCAGCAGGATATATCTCTCTGACGAAACTACGGCACTGCAGATGACTTGAAGCGC	420	
DB	251	GGATGAAGAACAATATACAAAATTAATGAAGACTTTGGTACTGCTGATGATTTGAAGTCTT	310	
QY	421	TCCTCTTCCGGCCCTTCATGAGAGGGGGATGTATCTTATGGTTCGATGTGGTTGCTTAACCATTA	480	
DB	311	TGGCACAGAATTTCCAGATCGTGATATGTTGTTAATGGTGGATATCGTTACCAACCATTT	370	
QY	481	TGGGCTATGATGAGCGGGTAGCTCAGTGATTAAGTGTTTAAACCGTTCAAGTTCCCTCCC	540	
DB	371	ACGGCAGTGTATGGCAGTGGAGATAGTATCGATTAATCTCAGAGTACACCCCGTTCAACGACC	430	
QY	541	AAGACTACTTCCACCCGCTTCGTGTTCAATCAAAACTATGAAGATCAGACTCAGGTTTCAGG	600	
DB	431	AAAAGTACTTCCATACTCTGCTTATTTTCAAACTATGATGACCAAGCTCAGGTTCAA	490	
QY	601	ATTGCTGGCTAGGAGATAACAATGTCTCTCTTGCTGTATCTCGATACCAAGGATGTGG	660	
DB	491	GTTCGCTGGGAAGGTGACTCTTTCAGTTCGATTAACAGATTTGAGAACGGAAGATAGCGACG	550	
QY	661	TCAAGAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTCAGCGGCC	720	
DB	551	TGGCCTCAGTTTCAATCTTTGGGTTAAAGATTTTGTGGCAATTACTCAATGTATGTT	610	
QY	721	TCCGTATCGACACAGTAAAAACAGTCAGGAAGGACTTCTGGCCCGGGTACAAACAAGCCG	780	
DB	611	TAAGAAATTGATGTGCTAAACATGTGACCAAGGCTTTTCCCGGATTTTGTGTAGTGCAT	670	
QY	781	CAGCGGTGATGTATCGCGGAGGTGCTCGACGGTGTATCGGCTCACTGTGTCCCTACC	840	
DB	671	CTCGAGTTTACTCAGTAGGCGGAAGTTTTCGAAGGAGACCCAGCTTATACATGCCCATACC	730	
QY	841	AGAAACGCTGAGCGGGTACTGAATCATCCCATTTACTTCACTCTCTCAAGCCCTTCA	900	
DB	731	AAATTTACATTCAGGGGTTAGTAAATTTATCCATTGTAATACCCNAACACAGATTTTTTA	790	
QY	901	AGTCAACCTCCGGCAGCATGGACGACTCTACAACTGATCAACACCGTCAAAATCCGACT	960	
DB	791	AAACTACTGATTCNAAGTTCAGTGTGAGTTGACTCAAAATGATTTTCAAGCGTTGCTTCCAGTT	850	
QY	961	GTCAGACTCAACACTCTCGGGCACAATTGCTCGAGAACACGACAAACCCACGGTTCGCTT	1020	
DB	851	GTTTCGATCCAACTTTGTTGACAAACTTTGTAGAAATACGATATGAAGGTTTCGCTT	910	
QY	1021	CTTACACCAACGACATAGCCCTCGCCGAAGAACGTCGAGCATTCATCATCTCTCAACGACG	1080	

RESULT 2

US-08-470-702-4

Sequence 4, Application US/08470702

Patent No. 5631149

GENERAL INFORMATION:

APPLICANT: MATSUI, IKUO

APPLICANT: ISHIKAWA, KAZUHIKO

APPLICANT: MIYAIRI, SACHIO

APPLICANT: HONDA, KOICHI

TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,

TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

STREET: 8110 Gatehouse Road, Suite 500 East

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,702

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/204,656

FILING DATE: 02-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: WEINER, MARC S.

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 234-252P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:


```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-467-831-4

Query Match      21.6%; Score 375.4; DB 1; Length 1404;
Best Local Similarity 55.6%; Pred. No. 2.9e-99;
Matches 752; Conservative 0; Mismatches 586; Indels 15; Gaps 1;

Qy 121 CTGGGACTGGCGATCGCAATCCATTTATTTCTTCTCAAGGATCGATTTGCAAGGACGG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11 CTGATAAATGGAGATCACAGTCATTTATCAAAATGTTCATGACAGATTTGCTAGAACCG 70

Qy 181 ATGGGTCGACGACTGCGACTGTGTAATACTGCGGATCAGAAATCTGTGGTGGAAACATGGC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 ATGGTGATACAAGTGCTTCCTGTAACACAGAAGATAGACTTTACTGTGGTGGTCTTTTCC 130

Qy 241 AGGCGATCATCGACAAGTTGGACTATATATCCAGGGAATGGGCTTCACAGCCATCTGGATCA 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 AAGGCATCAAAAGAAGTTGGATTACATCAAAAGATATGGGCTTTACTGCTATTTGGATTT 190

Qy 301 CCCCGCTTACAGCCAGCTGCCCCAGACACCGCATATGGAGATGCCCTACCATGGCTACT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 CTCAGTTGTTGAACAACATTTCCCGATTAACACAGCATATGGTTATGCTTATCTGTAACT 250

Qy 361 GGCAGCAGGATATATACTCTCTGAACGAAACCTACGGCAGCTGCAGATGACTTTGAAGCGC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251 GGATGAAGAACATATACAAAATTAATGAAACCTTTGGTACTGCTGATGATTTGGAAGTCT 310

Qy 421 TCTCTTCGGCCCTTCATGAGAGGGGAGTATCTTATGTGTCGATGTGGTGTGCTTAACCAT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 TGGCACAGAAGATTGACAGCATCGTGATATGTTGTTAAATGGTGGATATCGTTACCAACCAT 370

Qy 481 TGGGCTATGATGGAGCGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCC 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 ACGGCACTGATGGCAGTGGAGATAGTATCGAATTACTCAGAGTACACCCGGTTCAAGACC 430

Qy 541 AAGACTACTTCCACCCGTTCTGTTTTCATTCAAAACTATGAAGATCAGACTCAGGTGAGG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 AAAAGTACTTCCATAAATACTGCTCTTATTTCAAACCTATGATGACCAAGCTCAGGTTCAA 490

Qy 601 ATTGCTGGCTAGGAGATAAACACTGCTCTCTTGGCTGTATCTCGATACCACAGGATGTGG 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 GTTGCTGGGAAGGTGACTCTTTCACTTGCTATTAACAGATTTTGAGAACGGAAGATAGCA 550

Qy 661 TCAAGAATGAATGCTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCAATGCAGCGCC 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 TGGCCTCAGTTTTCAATCTTTGGGTTAAAGATTTTGTGGCAATTACTCAATGATGGTT 610

Qy 721 TCCGTATCGACACAGTAAAAACCGTCCAGAGGAGACTTCTGCGCCCGGGTCAACAAACCG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 611 TAAGAAATTGATAGTCTAAACATGTTGGACCAAGGCTTTTCCCGGATTTTGTAGTGCAT 670

Qy 781 CAGCGGTGTACTGTATCGCGGAGGTGCTCGACGGTGATTCGGGCTTACACTTGTCCCTACC 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 CTGGAGTTTACTCAGTAGGGCGAAGTTTTTCCAAGGAGACCCAGCTTATACATGCCCATACC 730

Qy 841 AGACGTCATGGACGGGCTACTGAATATCCATTTACTATCCACTCTCAACGCCCTTCA 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AAAATTATCATTTCCAGGGGTTAGTAATTAATTCATTTGTACTACCCCAACACGAGATTTT 790

Qy 901 AGTCAACCTCCGGCAGCATGGACGACCTCTCAACATGATCAACACCGTCAAAATCCGACT 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 AAATACTGATTCAAGTTCAGTGGATTGACTCAAAATGATTTCAAGCGGTGCTTCCAGTT 850

Qy 961 GTCCAGACTCAACACTCCTGGGCAATTTGTCGAGAAACCAACGACCAACCCACGGTTGCTT 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
;
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1404 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: other nucleic acid
;   DESCRIPTION: /desc = "Synthetic DNA"
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   IMMEDIATE SOURCE:
;   CLONE: Derived from plasmid pSF'1 (Agric. Biol. Chem.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1404
;   OTHER INFORMATION: /note= "Nucleotides 1-1404
;   OTHER INFORMATION: correspond to nucleotides 79-1482 in the Saccharomycopsis
;   OTHER INFORMATION: fibuligera '-amylase structural gene"
;
US-08-204-656B-1

Query Match      21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 8.3e-98;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 CTGCGGACTGGCGATCGCATCATTTATTTCTCTCAGGATCGATTGTCAGGACCG 180
DB 11 CTGATAATGGAGATCACAGTCTATTATCAAAATGTCTAGACAGATTGCTAGAACCG 70
QY 181 ATGGGTCGACGACTGGGACTTGTAAATCTCGGATCAGAAATCTGTGGTGGAACTGGC 240
DB 71 ATGGTGATACAGTGTCTCTGTAAACACAGAGATAGACTTACTGTGGTGTCTTTCC 130
QY 241 AGGGCATCATCGAAGTGGACTATATCCAGGGAATGGGCTTCAAGCAATCTGGATCA 300
DB 131 AAGGCATCATAAAGAGTGGATTACATCAAGATATGGGCTTTACTGCTATTGGATT 190
QY 301 CCCCGCTACAGCCGAGTCCCGACAGACCCGATATGAGATGCTTACCATTGGTACT 360
DB 191 CTCAGTGTGTAAACATTTCCCGATTAACACAGCATATGGTATGCTTATCATGGTTCT 250
QY 361 GGCAGCAGGATATATCTCTCTGAACGAAACTACGCGCATCGAGATGACTTGAAGCGC 420
DB 251 GGATGAAGACATATACAAATTAATGAACACTTTGGTACTGCTGATGATTGAGTCTT 310
QY 421 TCTCTTTGGCCCTTCATGAGAGGGGATGATCTTATGTCGATGTTGGTGTAAACATA 480
DB 311 TGGCAACAAGAAATGACAGTCTGTATGTTTAAATGGTGGATATCGTTTACCACCAAT 370
QY 481 TGGGCTATGATGAGCGGGTAGCTCAGTCGATTCAGTGTGTTTAAACGTTTCAAGTCCC 540
DB 371 ACGGCACTGATGCGAGTGGAGATAGTATCGATTACTCAGAGTACACCCCGTTCAACGACC 430
QY 541 AAGACTACTTCCACCGCTTCTGTTTCATTTCAAACTATGAAGATCAGATCAGGTTGAGG 600
DB 431 AAAGTACTTCCATACTACTGTCTTATTTCAAACTATGATGACCAAGCTCAGTTCAAA 490
QY 601 ATTGCTGGCTAGGAGATAACACTGTCTCTGCTGCTGATCTCGATACCAACCAAGGATGTG 660
DB 491 GTTGTGGGAAGTGAAGTCTTTCAGTTGCTTACAGATTTGAGAACCGAAGATAGCGAGC 550
QY 661 TCAGAAATGAATGGTACGAGTGGTGGGATCATTTGTTATCGAACTACTCATTTGACGGCC 720
DB 551 TGGCCTCAGTTTCAATTTCTTGGGTTTAAAGATTGTTGGCAATTTACTCAATTTGATGGTT 610
QY 721 TCCGTATCGACACAGTAAACACGTCAGAGAGACTTCTGGCCGGGTACAAACAGCCG 780
DB 611 TAAGAAATGATGCTTAAACATGTGGAACAGGCTTTTCCCGGATTTGTTAGTCCAT 670
QY 781 CAGGCGGTGATCTATCGGCGAGGTGCTCGACGGTGTATCGGCTTACACTTGTCCCTACC 840
DB 671 CTGGAGTTTACTCAGTAGGCGAGTTTTCACAGGAGACCCAGCTTATACATGCCATACC 730
QY 841 AGAACGTCTATGAGCGGCTACTGAACTATCCCAATTTACTATCCACTTCTCAACGCTTCA 900
```

```
DB 731 AAAATTACATTCAGGGGTTAGTAATTATTCATTGTACTACCAACCAAGAGATTTTTTA 790
QY 901 AGTCAACCTCCGGCAGCATGGACGACCTCTACAAACATGATCAACACCGTCAAAATCCGACT 960
DB 791 AAATCTACTGATTCAGGTTCCAGTGTAGTCTCAAAATGATTTCAAGCGTTGCTTCCAGTT 850
QY 961 GTCCAGACTCAACACTCTCTGGGACATTTGTCGAGAACCAACGACCAACCGTTCGCTT 1020
DB 851 GTTCGGATCCAACTTTGTTGACAAAATTTGTAGAAAATCAGGATAATGAAAGGTTGCTT 910
QY 1021 CTTTACACCAACGACATAGCCCTCGCCAAAGACGTGCGAGCATTCATCTCTCAAGGACG 1080
DB 911 CAATGACGAGGACCAAGTTTGAATTTCTAATGCTATTGATTTGCTTTTGGGTGATG 970
QY 1081 GAATCCCATCATCTACGCGGCCAAGAACAGACATCTACGCGCGGAAACGACCCGCGA 1140
DB 971 GTATTCTCTGCTATTACTATGACAAAGAACAGGCTTGAGCGGAAAAGTGACCCAAACA 1030
QY 1141 ACGCGAAGCAACCTGGCTCTCGGCTACCCGACGACGAGCTGTACAAAGTTAATTG 1200
DB 1031 ACAGAGAGGCTTGTGGTTATCCGCTACAAACAAAGAGAGTACTATTACAGCTCATTTG 1090
QY 1201 CTTCCGCGAACGCAATCCGGAATATGCTTAGCAAAAGATACAGGATTCGTGACCTTACA 1260
DB 1091 CCAAGCTAATGCTGCCAAGAACGCGCGCTTTATCAAGACTCAAGCTATGCCACCTCGC 1150
QY 1261 AGAATCGGCCCATCTCTACAAAGACGACACAAACGATCCCATGCGCAAGGCGACAGATGGT 1320
DB 1151 AGCTTTCTGTGATCTTTTCAATGACCATGTTATTGCAACAAAGAGCGAGCGTTGTTT 1210
QY 1321 CGCAGATCGTACTATCTTGTCAACAAAGGTCCTTCGGGTGATTCGTATACCTCTCCT 1380
DB 1211 C-----TGTTTTCAACAACTTTGGTTCCAGCGGTTCTTCTGATGTGACTA 1255
QY 1381 TGAGTGGTGGGTTACACAGCGCGCCGCAATTTGACGGAGGTCATTGGCTGCACGACCG 1440
DB 1256 TTTCCACACAGCTTACAGTTCCGGTGAGGATTTGGTAGAAGTTTTCACATGACGACTG 1315
QY 1441 TGACGGTGGTTCCGATGGAATTTGCTCTGTTTC 1473
DB 1316 TTAGCGGACGCTTGACTTTACAAGTTTCTATCC 1348
```

RESULT 5

```
US-08-204-656B-3
; Sequence 3, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
;   APPLICANT: Matsui, Ikuo
;   APPLICANT: Ishikawa, Kazuhiko
;   APPLICANT: Miyairi, Sachio
;   APPLICANT: Honda, Koichi
;   TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
;   TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
;   TITLE OF INVENTION: Oligosaccharide Using The Enzyme
;   NUMBER OF SEQUENCES: 18
;   CORRESPONDENCE ADDRESS:
;   ADDRESSES: Birch, Stewart, Kolaesch & Birch, LLP
;   STREET: 8110 Gatehouse Road, Suite 500 East
;   CITY: Falls Church
;   STATE: Virginia
;   COUNTRY: U.S.A.
;   ZIP: 22042
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent In Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/204,656B
;   FILING DATE: 02-MAR-1994
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
```

NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic nucleic acid"
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Derived from plasmid pSf'1 (Agric. Biol. Chem.
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1404
OTHER INFORMATION: /note = "Nucleotides 1-1404
correspond to nucleotides 79-1482 in the Saccharomycopsis
fibuligera '-amylase structural gene"
US-08-204-656B-3

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Mismatches 587; Indels 15; Gaps 1;
Matches 751; Conservative 0;

QY	121	CTGCGGACCTGGCGATCGCAATCATTTATTTCTCTCACCGATCGATTGCAAGACGG	180
DB	11	CTGATAAATGGAGATCACAGTCTTATTAATCAATTTCTCACTACAGATTGCTAGAACG	70
QY	181	ATGGGTGACGACTGCGACTTGTAACTGCGGATCAGAAATCTGTGTGGAACATGGC	240
DB	71	ATGGTGATACAAAGTGTCTCTGTAACACAGAGATAGACTTTTACTGTGTGTGTTCTTTCC	130
QY	241	AGGGATCATCAGACGATGGACTATATCCAGGATGGCTTACAGCCATCTGGATCA	300
DB	131	AAGGCATCAAGAAGTGGATTAATCATCAAGATATGGGCTTACTGCTATTGGATT	190
QY	301	CCCGGTTACAGCCAGCTGCCCGACACCGCATATGAGATGCTTACCATTGGCTACT	360
DB	191	CTCCAGTGTGAAACATTTCCGATTAACACAGCATATGTTATGCTTATCATGTTGGT	250
QY	361	GGCAGCAGATATATCTCTGAAACGAAATCTACGCGCATCGAGATGACTTTGAAAGGCG	420
DB	251	GGATGAAGAACATATACAAATTAATGAATACTTTGGTACTGCTGATGATTTGAAGTCTT	310
QY	421	TCTCTCGGCCCTTCATGAGAGGGGATGATCTTATGCTGATGTTGGTGTACCATTA	480
DB	311	TGGCAAGAAATGGCAGATCGATGATGTTGTTATGTTGGATATCGTTACCAACCAAT	370
QY	481	TGGGCTATCATGGAGCGGGTAGCTCAGTCGATTAACAGTGTGTTAAACCGTTTCA	540
DB	371	ACGGCAGTGTGGCAGTGGAGATGATGATGATTAATCAGAGTACACCCCGTTCAAGGAC	430
QY	541	AAGACTACTTCCACCGGTTCTGTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGG	600
DB	431	AAAAGTACTTCCAACTACTGCTCTTATTTCAAACTATGATGACCAAGCTCAGGTTCAA	490
QY	601	ATTGCTGGCTAGAGATAACATGCTCTCTGCTGATCTCGATACCACCAAGGATGGG	660
DB	491	GTTGCTGGGAAGGTGACTCTTTCAGTTGCAATACAGATTTGAGAACGGAGATAGCGAG	550
QY	661	TCAAGAAATCAATGGTACGACTGGGTGGGATCATTTGTTATCGAACTACTCATTGACGGCC	720
DB	551	TGGCTCAGTTTCAATTTCTGGTTAAAGATTTGTTGGCAATTAATCAATGATGGTT	610
QY	721	TCCGTATCAGACAGTAAACACAGTCCAGAGGACTTTGGCCCGGGTACAAACAAAGCGC	780

DB	611	TAAGAAATTTAGTAGTCTAAACATGTGGACCAAGGCTTTTTCGGGATTTTGTAGTGCAT	670
QY	781	CAGGCGTGTACTGTATCGCGAGGCTCGACGGGTATCCGGCCCTACACTGTCCCTACC	840
DB	671	CTGGAGTTTACTCAGTAGCGGAAGTTTCCAAAGGAGACCCAGCTTATATGCCCATACC	730
QY	841	AGAACGTCTATGGACGGCGTACTGAACTATCCCATTTACTTATCTCTCAACGCTTCA	900
DB	731	AAAATTACATTCACGGGGTTAGTAATTAATTCATTCTACTTACCACCAACGAGATTTT	790
QY	901	AGTCAACCTCCGGCAGCATGGACGACCTCTCAACATGATCAACCGTCAAAATCGGACT	960
DB	791	AAACTACTGATTCAAGTTTCCAGTGAGTTGACTCAAAATGATTTCAAGCGTGTGCTCCAG	850
QY	961	GTCCAGACTCAACACTCTCTGGGCATCTCTGTCGAAACACGACCAACCGGTTGCTTT	1020
DB	851	GTTGGGATCCAACTTTGTTGACAAACTTTGTAGAAAATCAGATAATGAAAGTTTGGCT	910
QY	1021	CTTACACCAACGACATAGCCCTCGCCAAAGACGTCGACAGCTTTCATCTCTCAACGACG	1080
DB	911	CAATGACCCAGCGACCAAGTTTGATTCTTAATGCTATTGCAATTTGCTCTTTTGGGTGATG	970
QY	1081	GAAATCCCATATCTACCGCGGCCAAGAACAGCACTACGCGCGGAAACGACCCGCGA	1140
DB	971	GTAATCTCTGCTATTTACTATGACAAAGAACAGGCTTGAGCGGAAAGAGTGACCCAAACA	1030
QY	1141	ACCGGNAAGCAACCTGGCTCTCGGGCTACCCGACGACGAGCGAGCTGTACAAAGTTAATG	1200
DB	1031	ACAGAGAGCGCTTGTGTTATCCGGCTACAAACAAAGAGAGTGACTATTACAAGCTCATTTG	1090
QY	1201	CCTCCGGAAGCAATCGGAACTATGCCATTAGCAAAAGATACAGGATTCGTGACCTACA	1260
DB	1091	CCAAAGCTAATGCTCCAGAAACGCGCGTTTATCAAGACTCAAGCTATGCCACCTCGC	1150
QY	1261	AGAACTGGCCCATCTCAAAAGACGACACAAAGATCGCATCGCAAGGCAAGGCAAGATGGT	1320
DB	1151	AGCTTTCTGTGATCTTTTCAAAATGACCATGTTATTGCAACAAAGAGAGCGAGCTGTTT	1210
QY	1321	CGCAGATCGTACTATCTTGTCCAAAGGGTGTCTCGGGTATTCGTATATACCTCTCCT	1380
DB	1211	C-----TGTTTTCAACAACCTTGGTTCAGCGGTTCTTCTGATGTGACTA	1255
QY	1381	TGAGTGTGCGGGTTACACAGCCGCGCAATTCAGCGAGGTCAATGGCTGCACGACG	1440
DB	1256	TTTCCAACACAGGTTACAGTTCGCGTGAGGATTTGTTAGTAAGTTTTGACATCAGTACTG	1315
QY	1441	TGACGGTGTGTCGATGGAATGTCCTGTTTTC	1473
DB	1316	TTAGCGGAGCTCTGACTTACAAAGTTTCTATCC	1348

RESULT 6

US-08-204-656B-5
; Sequence 5, Application US/08204656B
; Patent No. 553882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuro
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1404 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "Synthetic DNA"

HYPOTHETICAL: NO

ANTI-SENSE: NO

IMMEDIATE SOURCE:

CLONE: Derived from plasmid pSF1 (Agric. Biol. Chem.

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1404

OTHER INFORMATION: /note= "Nucleotides 1-1404

OTHER INFORMATION: correspond to nucleotides 79-1482 of the Saccharomycopsis

OTHER INFORMATION: fibuligera 'amylase structural gene"

US-08-204-656B-5

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 8.3e-98;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;
QY 121 CTGCGGACTGGCGATCGCAATCATTTATTCCTCTCTCAGGATCGATTGTCGAAGACGG 180
DB 11 CTGATAATGGAGATCACAGTCTATTATCAATTTGCTCACTGACAGATTGCTGAGAACC 70
QY 181 ATGGGTGCGAGCTGGGACTGTTGTAATCTCGGATCAGAAATATCTGTGTGGAACTGGC 240
DB 71 ATGGTGATACAGTGTCTCTGTAACACAGAGATAGACTTTACTGTGTGGTCTTTTC 130
QY 241 AGGGCATATCGAAGTTGGACTATATCCAGGATGGCTTCAGACCATCTGGATCA 300
DB 131 AGGGCATATAAAGAGTTGGATTACATCAAGATATGGGCTTTACTGTCTATTGGATT 190
QY 301 CCCCCGTTACAGCCAGCTGCCCGACACACCCGATATGGAGATGCTTACCATGGCTACT 360
DB 191 CTCAGTTGTTGAACATTTCCGATACACAGCATATGGTTATGCTTATCATGGTCTCT 250
QY 361 GGCGAGGATATATATCTCTCTGAACGAAATCTACCGCATCGAGATGATCTGAAGGCGC 420
DB 251 GGATGAAGAACATATACAAATTAATCAAACTTTGGTACTGTCTGATGATTTCAGTCTT 310
QY 421 TCTCTTCGGCCCTTCATGAGGGGGATGATCTTATGGTCGATGTGGTCTTAACCATTA 480
DB 311 TGGCAACAAGATTCGACGATCGTGATATGTTGTAATGGGATATGCTTACCAACCAT 370
QY 481 TGGGCTATGATGAGGGGGTAGCTCAGTTCGATTACAGTGTGTTTAAACCGTTCAGTTCCC 540
DB 371 ACGGCAAGTATGCGAGTGGAGATAGTATCGATTACTCAGAGTACACCCGTTCAACGCC 430
QY 541 AGAGTACTTCCACCGCTTCTGTTTCATTCAAACTATGAAGATCAGACTCAGGTTGAGG 600
DB 431 AAAAGTACTTCCATACTACTGCTTATTATTAATCAAACTATGATGATCAAGCTCAGGTTCAA 490
QY 601 ATTGCTGGTCAAGATACACTGTCTCTCTGCTGATCTCGATACCAACCAAGATGTGG 660

DB 491 GTTGTGGGAAGGTGACTCTCTCAGTTGTCATTACAGATTGAGAACGGAAGATAGCGACG 550
QY 661 TCAGAAATGAATGGTACGACTGGTGGGATCATTTGGTATCGAATCTACTCCATTGACGGCC 720
DB 551 TGGCCTCAGTTTTCATTTCTGGGTTAAAGATTTTGTGGCAATTAATCAATTCATGATGGTT 610
QY 721 TCGGTATGACACAGTAAACACGCTCCAGAGGACTTCTGGCCGGGTACAAACAAAGCCG 780
DB 611 TAAGAAATGATAGTCTAAACATGTGGACCAAGGCTTTTCCGGATTTTGTGTAGTCAT 670
QY 781 CAGGGGTGATCTGATCGCGAGGTGCTCGACGGTGTATCGGCTTACACTTGTTCCTTACC 840
DB 671 CTGGAGTTTACTCAGTAGGCGAAGTTTCCAAAGGAGACCCAGCTTATACATGCCCATACC 730
QY 841 AGAACGTATGAGACGGCGTACTGAATATCCATTTACTATCCATCTCTCAACGCCCTTCA 900
DB 731 AAAATTTACATTCAGGGGTTAGTAATTTATCTCAATTTGTACTACCCCAACACGAGATTTT 790
QY 901 AGTCAACCTCCGCGACATGGACGACCTCTACAAATGATCAACACCGTCAAAATCCGACT 960
DB 791 AAATCTAGTTCAAGTTCCAGTGTGACTCAAAATGATTTCAAGCGTTGCTTCAGTT 850
QY 961 GTCCGAGCTCAACACTCTCTGGGCACATTCGTGAGAACCCAGCAACCCAGCGTTCCGCTT 1020
DB 851 GTTCGGATCCAACTTTGTTGACAACTTTGTAGAAAATCAGGATAATGAAAGGTTCCGCTT 910
QY 1021 CTTACACCAACGACATAGCCCTCGCCCAAGACGTGCGACGATTCATCTCTCAACGACG 1080
DB 911 CAATGACGACGACCAAAAGTTTGAATTTCTAATGCTATTGCTCTTTTGGGTGATG 970
QY 1081 GAATCCCATCATCTACGCGCGCCGAGACAGCAGCTACGCGCGGGAAGACCCCGCA 1140
DB 971 GTATTCTCTGCTATTACTATGAGCAAGAACAAAGGCTTGAGCGGAAAAGTGACCCCAACA 1030
QY 1141 ACCGGAAGCAACCTGGGCTCTCGGCTACCCGACGACGAGCTGTACAACTTAATTG 1200
DB 1031 ACAGAGAGCGCTTGTTGTTATCGGCTACACAAAGAGAGTACTATTACAAGCTATTG 1090
QY 1201 CTTCCGGAACCGAATTCGGAATCTATGCAATTTAGCAAAAGATACAGGATTCGTGACCTACA 1260
DB 1091 CCAAGCTAATGCTGCGAAGAACGCGCGCTTTATCAAGACTCAAGCTATGCCACCTCGC 1150
QY 1261 AGNACTGGCCCATCTACAAAGACGACACAAAGATGCCATGCCAGGCGACAGATGGGT 1320
DB 1151 AGCTTTCTGTGATCTTTTCAATGACCATGTTATTGCAACAAAGAGGAGCGGTTGTT 1210
QY 1321 CGCAGATCGTACTATCTTGTCCAAAGGGTGTCTCGGCTGATTCGTATACCTCTCCT 1380
DB 1211 C-----TGTTTCAACAACTTGGTTCAGCGGTTCTTCTGATGACTA 1255
QY 1381 TGAGTGGTGGGTTTACACAGCGCGCCAGCAATTCAGGAGGTCAATGGCTGCACACCG 1440
DB 1256 TTTTCAACACAGGTTACAGTTCGCGTGAGGATTTGGTAGAAGTTTGTGACATGCACTG 1315
QY 1441 TCACGGTTGGTTCGATGGAATGTGCTGTTTC 1473
DB 1316 TTAGCGGACGCTGACTTTACAAGTTTCTATCC 1348

RESULT 7

US-08-470-702-1

; Sequence 1, Application US/08470702

; Patent No. 5631149

; GENERAL INFORMATION:

; APPLICANT: MATSUI, IKUO

; APPLICANT: ISHIKAWA, KAZUHIKO

; APPLICANT: MIYAKI, SACHIO

; APPLICANT: HONDA, KOICHI

; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,

; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-1

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 8.3e-98;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 CTGCGGACTGGCGATCGCAATCATTTATTTCTCTCAGGATCGATTGTCGAAAGCGG 180
DB 11 CTGATAAATGGAGATCACAGTCTATTATCAAAATGTCTACAGACAGATTGCTAGAACGG 70
QY 181 ATGGGTGCGAGCTGCGACTTGTATATCTCGCGATCAGAAATACTGTGTGGAATGGC 240
DB 71 ATGGGTGATCAAGTGTCTCTGTAACACAGAAAGATAGACTTTTACTGTGTGTGTTCTTCC 130
QY 241 AGGGCATCATGACAAAGTTGGACTATATCCAGGGAATGGCTTTCAGGCCATCTGGATCA 300
DB 131 AAGGCATCATAAAGATTGGATTATCATCAAGATATGGGCTTTACTGTATTGGATT 190
QY 301 CCCCCGTACAGCCAGCTGCCAGACACCCGCATATGAGATGCTTACCATGGCTACT 360
DB 191 CTCCAGTTGTGAAACATATCCCGATAACACAGCATATGTTATGTTATCTATCATGTTTCT 250
QY 361 GGCAGCAGATATATATCTCTGACGAGAAACTACGGCACTCAGATGACTTTGAAGGCGC 420
DB 251 GGATGAAGAACATATACAAATTAATGAAGAACTTTGGTACTGCTGATGATTTGAAGTCTT 310
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGTTGATGTTGTTGCTAAACCAT 480
DB 311 TGGCACAGAAATGACAGATCGTATATGTTTAAATGTTGATATCGTTACCAACCAT 370
QY 481 TGGGCTATGATGAGCGGGTAGCTCAGTCGATTACAGTGTGTTAAACCGTTTCAAGTCC 540
DB 371 ACGGCAGTGTGCGAGTGGAGATAGTATCGATTACTCAGAGTACACCCCGTTCAAGACC 430
QY 541 AAGACTACTTCCACCGGTTCTGTTTCATTTCAAACTATGAAGATCAGACTCAGTTGAGG 600
DB 431 AAAAGTACTTCCATACTACTGTTCTTATTTCAAACTATGATGACCAAGCTCAGGTTCAA 490

QY 601 ATTGCTGCTAGGAGATAAAGTCTCTCTTGGCTGATCTCGATACCAACCAAGGATGTGG 660
DB 491 GTTGTGGGAAGGTGACTCTTTCAGTTGCTATACCAAGATTGAGAAACGGAAGATAGCGACG 550
QY 661 TCAGAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAATCTACTCCATTGAGCGCC 720
DB 551 TGGCCTCAGTTTTCAAATTTCTTGGGTTAAAGATTTTGTGGCAATTAATCAATTTGATGTT 610
QY 721 TCCGTATCGACACAGTAAACACAGTCCAGAGGACTTCTGGCCCGGGTACAAACAAAGCGG 780
DB 611 TAAGAATTTGATAGTCTTAAACATGTGGACCAAGGCTTTTCCCGGATTTTGTGATGCTCAT 670
QY 781 CAGCGCTGTACTGTATCGCGAGGTGCTCGACGGTATCGCGCTTACACTTGTCCCTACC 840
DB 671 CTGGAGTTTACTCAGTAGGCGAAGTTTCCAAAGGAGACCCAGCTTATATCATGCCATACC 730
QY 841 AGAAGCTCATGACGCGGCTACTGAATCTATCCATTTTACTATCCACTCTCAAGCGCTTCA 900
DB 731 AAAATTACATTTCCAGGGGTAGTAAATTTATCTACTTACCCCAACCAACGAGATTTTATA 790
QY 901 AGTCAACCTCCGGCAGCATGGACGACCTCTCAACATGATCAACACCGTCAAAATCCGACT 960
DB 791 AAATCTGATTTCAAGTTCAGTGAGTTGACTCAATGATTTCAAGCGTTGCTTCCAGTT 850
QY 961 GTCCAGACTCAACACTCTCTGGGCACATTTGTCGAGAACCAACCAACCGTTCGCTT 1020
DB 851 GTTCGGATCCAACTTTGTTGACAAACTTTGTAGAAAATCAGCATTAATGAAGGTTGCGTT 910
QY 1021 CTTACACCAACGACATAGCCCTCGCCAAAGAGCTGCGAGCATTCATCTCTCAAGCAGC 1080
DB 911 CAATGACGAGCGACCAAAAGTTTGAATTTCTTAATGCTATTGCTATTTGCTCTTTTGGGTGATG 970
QY 1081 GAATCCCATCATCTACGCGCCCAAGAACAGCAGCTACGCGCGGGAACACGACCCGCGA 1140
DB 971 GTATTCCTGTCTATTACTATGGAACAAGAACAGGCTTGAGCGGAAAAGTGAACCAACA 1030
QY 1141 ACCGGAAGCAACCTGGCTCTCGGGCTACCCGACCGACGAGCTGTACAAAGTTAAATTG 1200
DB 1031 ACAGAGAGCGCTTGTGGTTATCCGGCTACCAACAAAGAGAGTGACTATTACAAGCTCATTTG 1090
QY 1201 CTTCCGGAACGCAATTCGGAATCTATGCCATTAGCAAAAGATACAGGATTCGTGACCTACA 1260
DB 1091 CCAAGAGCTAATGCTGCCAGAACGCGCGCTTTATCAAGACTCAAGCTATGCCCTCGC 1150
QY 1261 AGAATCGGCCCATCTACAAAGACGACACAAACGATCGCATCGCAAGGCGACAGATGGGT 1320
DB 1151 AGCTTCTGTGATCTTTCAAAATGACCATGTTATTGCAAAACAAAGAGCAGCGTTGTTT 1210
QY 1321 CGCAGATCGTGAATCTTTGTCACAAAGGGTGTCTTGGGGTGAATTCGTATACCTCTCTCT 1380
DB 1211 C-----TGTTTTCAACAACTTTGGTTCCAGCGGTTCTTCTGATGTGACTA 1255
QY 1381 TGAGTGTGGGTTTACACAGCGCGCCAGCAATTCAGCGAGGTCTATGGCTGACGACCG 1440
DB 1256 TTTCCAACACAGGTTTACAGTTCCGGTGAGGATTTGTTAGAAAGTTTGTGACATGCACTG 1315
QY 1441 TGACGTTGTTTCCGATGGAATGTGCTGTTTC 1473
DB 1316 TTAGCGGAGCTCTGACTTACAAGTTTCTATCC 1348

RESULT 8

US-08-470-702-2
; Sequence 2, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
S-08-470-702-2

Query Match	21.6%;	Score 373.8;	DB 1;	Length 1404;
Best Local Similarity	55.5%;	Prod. No. 8.3e-98;		
Matches 751;	Conservative 0;	Mismatches 587;	Indels 15;	Gaps 1;
121	CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTCACGGATCGATTTCGAAGCGG	180		
11	CTGATAAATGGAGATCAAGCTCTATTTATCAAATTTGTCACTGCACAGATTTCGTAAACCG	70		
181	ATGGGTCGACGACTCGGACTTGTAACTACTGGCGATCAGAAATACTGTGGTGGAAACATGGC	240		
71	ATGGTGATACAAGTGCTTCTTGTAACACAGAAGATAGACTTTACTGTGCTGTTCTTTC	130		
241	AGGGCATCATCGAAGAATGGACTATATCCAGGGAATGGGCTTCACAGCCATCTGGATCA	300		
131	AAGGCATCATAAAGAAGTTGGATTACATCAAAGATATGGGCTTTACTCTATTTGGATTT	190		
301	CCCCGGTTACGCCAGCTGCCCCAGACACCGCATATGGAGATGSCCTACCATGSCCTACT	360		
191	CTCCAGTTGTTGAAACATTTCCCGATACACACGATATGGTTATGCTTATCATGTTGGT	250		
361	GGCAGCAGGATATATCTCTTGTAACGAAAACCTACGGCACTGCGACATGTCAGATGACTTGAAGCGC	420		
251	GGATGAAGAACAATATACAAAATTAATGAAAACCTTTGGTACTGCTGATGATTTGAAGTCTT	310		
421	TCCTTCGCGCCCTTCATGAGAGGGGATGTATCTTTATGTCGTGATGTTGGTTCGTAACCAATA	480		
311	TGGCACAAGAAATGGACGATCGTGATATGTTTAAATGGTGGATATCGTTTACCAACCAATT	370		
481	TGGGCTATGATGGAGCGGGTAGCTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCC	540		
371	ACGGCAGTGATGGCAGTGGAGATGATGCTGATTTACTCAGAGTACACCCCGTTTCAAGACC	430		
541	AAGACTACTTCCACCCGTTCTGTTTCATTTCAAAACATATGAAGATCAGACTCAGGTTGAGG	600		

RESULT 9
US-08-470-702-3
: Sequence 3, Application US/08470702
: Patent No. 5631149
: GENERAL INFORMATION:
: APPLICANT: MATSUI, IKUO
: APPLICANT: ISHIKAWA, KAZUHIKO
: APPLICANT: MIYAIRI, SACHIO
: APPLICANT: HONDA, KOICHI
: TITLE OF INVENTION: VARIANT-TYPE

[illegible]

;; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
;; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
;; NUMBER OF SEQUENCES: 17
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
;; STREET: 8110 Gatehouse Road, Suite 500 East
;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22042
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/470,702
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/204,656
;; FILING DATE: 02-MAR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WEINER, MARC S.
;; REGISTRATION NUMBER: 32,181
;; REFERENCE/DOCKET NUMBER: 234-252P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 205-8000
;; TELEFAX: (703) 205-8050
;; TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1404 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (synthetic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-470-702-3

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 8.3e-98;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY	121	CTGCGGACTGGCGATCGCAATCAATTATTCTCTCAGCGATCGATTGCAAGACGG	180
DB	11	CTGATAAATGGAGATCACAGTCTATTATCAAAATGCTCACTGACAGATTGCTAGA	70
QY	181	ATGGGTCGACGACTGCGACTTGTATATCTGCGGATCAGAAATACTGTGGAAATGGC	240
DB	71	ATGGTGATACAGTGTCTCTGTACACAGAGATAGACTTACTGTGTGGTCTTTCC	130
QY	241	AGGGCATCATCGCAAGTTGGACTATATCCAGGAATGGGCTTACAGCCATCTGGATCA	300
DB	131	AAGGCATCATAAAGATTGGATTACATCAAGATATGGGCTTACTGCTATTGGATT	190
QY	301	CCCCGTTACAGCCAGTGCCTCCACACACCGCATATGAGATGCTACCATGGTACT	360
DB	191	CTCCAGTTGTTGAAACATTTCCCGATTAACACAGCATATGTTATGCTTATCATG	250
QY	361	GGCAGCAGGATATATACTCTCTGAAACGAAACTPACGGCACTGCAGATGACTTGA	420
DB	251	GGATGAAGAACATATACAAATTAATGAACACTTTGGTACTGCTGATGATTTGA	310
QY	421	TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATATGGTCGATGGTTGCTAAC	480
DB	311	TGSCAACAAAGATTGACAGTCTGTGATATGTGTTAATGTGGATATCGTTACCA	370
QY	481	TGGGCTATGATGGAGCGGGTAGCTCAGTCGATCAGTGTGTTAAACCGTTCA	540
DB	371	ACGGCATGATGGCAGTGGAGATAGTATCGATTACTCAGAGTACACCCCGTTCA	430

QY	541	AAGACTACTTCCACCCGTTCTGTTTCATTCAAAACATATGAAGATCAGACTCAGGTTGAGG	600
DB	431	AAAAGTACTTCCATAAATACTGTCTTATTTCAAAACATATGAATGACCAAGCTCAGGTTCAA	490
QY	601	ATTGCTGCTAGGAGATAAATACTGTCTCTTGGCCCTGATCTCGATACCAACCAAGGATGGG	660
DB	491	GTTGCTGGGAAGGTGACTCTTCAGTTGCAATTACAGATTGGAACGGAGATAGGAGC	550
QY	661	TCAAGATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGCC	720
DB	551	TGGCCTCAGTTTCAATTTCTTGGGTTAAAGATTTTGTGGCAATTACTCAATTGATGGTT	610
QY	721	TCCGTATCGACACAGTAAAAACACGTCCAGAGGACTTCTGGCCCGGGTACAACAAGCGG	780
DB	611	TAAGAAATTCATAGTCTAAACATGTGGACCAAGGCTTTTCCCGGATTTTGTAGTCAT	670
QY	781	CAGGCGTGTACTGTATCGCGAGGTGCTCGAGGTGATCCGGCTACACTTCTCCCTACC	840
DB	671	CTGGAGTTTACTCAGTAGGCGAAGTTTCCAAAGGAGACCCAGCTTATCATGCCATACC	730
QY	841	AGAACGTCATGGACGGCGTACTGAACTATCCCACTTACTTATCTCCACTCTCCACGGCTTCA	900
DB	731	AAATTTACATTCACGGGTTAGTATATTATCCATTCTACTACCCACACGAGATTTTTTA	790
QY	901	AGTCAACCTCCGGCAGCATGAGACCTCTCAACATGATCAACCGTCAAAATCCGACT	960
DB	791	AAACTACTGATTCAAAGTTCAGTGAGTTGACTCAAAATGATTTCAAGCGTTGCTCCAGTT	850
QY	961	GTCCAGACTCAACACTCTCGGGCACAATTCGTGAGAACACGACACACCCAGGTTGCGTT	1020
DB	851	GTTGCGGATCCAACTTTGTGACAAAATTTGTAGAAAATCACGATATGAAAGGTTGCGTT	910
QY	1021	CTTACACCAACGACATAGCCCTCGGCAAGAACGTCGACGACTTCACTCTCAACGACG	1080
DB	911	CAATGACACGACGACCAAGTTGATTTCTAATGCTATTGCAITTTGCTCTTTGGGTGATG	970
QY	1081	GAATCCCATCATCTACGCCCGGCCAAGAACAGCACTACGCCGGCGGAAACGCCCGCGA	1140
DB	971	GTATTCTCTGTCAATTTACTATGACAAAGAACAAAGGCTTGAGCGGAAAAAGTGACCCAAACA	1030
QY	1141	ACCGGAAGCAACCTGGCTCTCGGGCTACCCGACGACGCGAGCTGTACAAGTTAATTG	1200
DB	1031	ACAGAGAGCGCTTGTGGTTATCCGGCTACAAACAAAGAGAGTACTATTACAAGCTCAITG	1090
QY	1201	CCTCCGCAACCAATCCGGAATATGCCAATAGCAAAAGATACAGGATTCGTGACCTACA	1260
DB	1091	CCAAAGCTAATGCTCCAGAAACGCCGCGTTTATCAAGACTCAAGCTATGCCACCTCGC	1150
QY	1261	AGAACTGGCCCCATCTACAAAGACGACAAACGATCGCCATGCGCAAGGCGACAGATGGGT	1320
DB	1151	AGCTTTCTGTGATCTTTTCAAAATGACCATGTTATTGCAACAAAAAGAGCGCGTTGTTT	1210
QY	1321	CGCAGATCGTACTATCTTGTCACAAAGGTGCTTCGGGTGATTCGTATACCTCTCCT	1380
DB	1211	C-----TGTTTTCAACAACTTGGTTCCAGCGGTTCTTCTGATGTGACTA	1255
QY	1381	TGAGTGGTGGGTTTACACAGCGGCCAGCAATTGACGGAGGTCAATTGCTGCACGACCG	1440
DB	1256	TTTCCAAACACAGGTTACAGTTCGGGTGAGGATTTGTAGAAGTTTGTACATGACGACTG	1315
QY	1441	TGACGGTTGGTTCCGATGGAATGTGCCTGTTTC	1473
DB	1316	TTAGCGGACGCTCTGACTTACAAGTTTCTATCC	1348

RESULT 10
US-08-467-831-1
; Sequence 1, Application US/08467831
; Patent No. 5635378
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO

APPLICANT: HONDA, KOICHI
 TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
 TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
 TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 STREET: 8110 Gatehouse Road, Suite 500 East
 CITY: Falls Church
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22042
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,831
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/204,656
 FILING DATE: 02-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: WEINER, MARC S.
 REGISTRATION NUMBER: 32,181
 REFERENCE/DOCKET NUMBER: 234-252P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 TELEX: 248345
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1404 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (synthetic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-467-831-1

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
 Best Local Similarity 55.5%; Pred. No. 8.3e-98;
 Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY	121	CTGCGGACTGGCGATCGCAATCCATTATTTCCTCTCAGGATCGATTGCAAGACGG	180
DB	11	CTGATAAATGGAGATCACAGTCTATTATCAAAATTGTCACTGACAGATTGCTAGAACCG	70
QY	181	ATGGGTGCGACGCTGCGACTTGTAACTACGCGATCAGAAATACTGTGGTGGAAACATGGC	240
DB	71	ATGGTGATACAGTGTCTCTGTAAACAGAAATAGACTTTACTGTGGTGTCTTTCC	130
QY	241	AGGGCATCATCGAAGTGGACTATATCCAGGAATGGGCTTTCAGACCATCTGATCA	300
DB	131	AAGGCATCAAAAGAGTGGATTATACATCAAGATATGGGCTTACTGCTATTGGATT	190
QY	301	CCCCCGTTACAGCCAGCTGCCAGACCCCATATGGAGATGCTTACCATTGGTACT	360
DB	191	CTCCAGTGTGTGAAACATTTCCCGATAACACAGCATATGGTTATGTCTATCATGGTTCT	250
QY	361	GGCAGCAGGATATATCTCTCTGAACGAAACTACGCGACTGCAGATGACTTGAAGCGC	420
DB	251	GGATGAAGACATATACAAATTAATGAACACTTTGGTACTGCTGATGATTGAGTCTT	310
QY	421	TCTCTTTCGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGTGGTTGCTAACCAT	480
DB	311	TGGCACAAAGAAATGACAGTCGTGATATGTTGTTAATGGTGGATATCGTTTACCACCAT	370
QY	481	TGGGCTATGATGGAGGGGTAGCTCAGTTACAGTGTGTTTAAACCGTTCAAGTCC	540

DB	371	ACGGCAGTGTGGCAGTGGAGATAGTATCGATTACTCAGAGTACACCCCGTTCAACGACC	430
QY	541	AAGACTACTTTCACCCGTTCTGTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG	600
DB	431	AAAAGTACTTCCATAACTACTGTCTTATTTCAAACCTATGATGACCAAGCTCAGGTTCAAA	490
QY	601	ATTGCTGGCTAGGAGATACACTGTCTCCTCGCTGATCTCGATACCAACCAAGGATGTGG	660
DB	491	GTGTCTGGGAAGGTGACTCTTTCAGTTGCAATACCAGATTTGAGAACGGAAGATAGGACG	550
QY	661	TCAAGAAATGAATGGTACGACTGGGTGGGATCAATTTGGTATCGAACTACTTCAATTGAGCGCC	720
DB	551	TGGCCTCAGTTTCAATTCTTGGGTTAAAGATTTTGTTGGCAATTACTCAATTGATGTT	610
QY	721	TCCGTATCGACACAGTAAACACGTCACGAAGACTTCTGGCCCGGGTACAAACAAAGCCG	780
DB	611	TAAAGAAATGATAGTGTAAACATGTGGACCAAGGCTTTTCCGGGATTTTGTAGTCCAT	670
QY	781	CAGGGGTGACTGTATCGGGGAGGTGCTCGAGGTGATCGGCGCTACACTTGTCCCTACC	840
DB	671	CTGGAGTTTACTCAGTAGGAGGAGTTTCCAAAGGAGACCCAGCTTATACATGCCCATACC	730
QY	841	AGAACGTCAATGGACGGCGTACTGAACTATCCCAATTTACTATCCACTCTCCAACGCTTCA	900
DB	731	AAATTTACATTCAGGGGTTAGTAAATTTATTCATTGTACTACCAACACGAGATTTTAA	790
QY	901	AGTCAACCTCCGGCAGCATGAGACCTCTCAACATGATCAACACCGTCAAAATCCGACT	960
DB	791	AAACTACTGATTCAGTTCCAGTGTGACTCAAAATGATTTCAAGCGTGTCTTCCAGTT	850
QY	961	GTCCAGACTCAACACTCTCTGGGACATTCGTGAGAACACGACCAACCCAGGTTGCTT	1020
DB	851	GTTCGGATCCAACTTTGTTGCAAACTTTGTAGAAATCAGGATAATGAAGGTTGCTT	910
QY	1021	CTTACACCAACGACATAGCCCTCGCCCAAGAACTGCGACGATTCATCATCTCAACGACG	1080
DB	911	CAATGACGAGGACCAAGTTTGATTTCTAATGCTATTGCTATTTGCTTTTGGGTGATG	970
QY	1081	GAATCCCCATCATCTACGCCGCCAAGAACAGACACTACCGCGCGGAAACGACCCCGCA	1140
DB	971	GTATTCCTGTCTTACTATGACAAAGAACAAAGGCTTGAGCGGAAAGTGACCCAAACA	1030
QY	1141	ACGGGAAGCAACCTGGCTCTCGGCTACCGACCGACGAGCGCTGTACAAGTTAATTG	1200
DB	1031	ACAGAGAGGCTTGTGGTTATCCGGCTACACAAAGAGAGTGACTATTACAAGCTCATTG	1090
QY	1201	CCTCGCGAAGCAATCCGGAACCTATGCAATTAGCAAAAGATACAGGATTCGTGACCTACA	1260
DB	1091	CAAAAGTAAATGCTGCAGAAACGCGCGCTTTATCAAGACTCAAGCTATGCCACCTCGC	1150
QY	1261	AGAACTGGCCCATCTACAAGACGACACAAACGATCGCCATGCGCAAGGGCACAGATGGGT	1320
DB	1151	AGCTTCTGTGATCTTTTCAATGACCATGTTATTGCAACAAAGAGCGAGCGTTGTTT	1210
QY	1321	CGAGATCTGATCATCTTGTCCAAACAGGGTGTCTCGGGTGATTCGTATACCTCTCCT	1380
DB	1211	C-----TGTTTTCAACAACTTGGTTCCAGCGTCTTCTCTGATGTGACTA	1255
QY	1381	TGAGTGGTTCGGGTATACAGCGCGCCCACTTGAAGGAGGTCAATTGGCTGCACGACCG	1440
DB	1256	TTTCCAACACAGGTTACAGTTCCGGTGAGGATTTGGTAGAAGTTTGCATGACGACTG	1315
QY	1441	TGACGGTGGTTCCGATGGAATGTGCTGTTTC	1473
DB	1316	TTAGCGGACGCTCTGACTTACAAGTTTCTATCC	1348

RESULT 11

US-08-467-831-2

; Sequence 2, Application US/08467831

; Patent No. 5635378

; GENERAL INFORMATION:

; APPLICANT: MATSUI, IKUO

APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE.
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-467-831-2

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 8.3e-98;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 CTGGGACTGGCGATCGCAATCCATTATTTCTCTCACGGATCGATTTGCAAGGACGG 180
DB 11 CTGATAAATGGAGATCACAGTCTAATTTATCAAAATGTCTACTACAGATTTGCTAGAACG 70
QY 181 ATGGGTGCGACGCTGCGATTTGTAATCTGCGGATCAGAAATCTGTGTGGAAACATGGC 240
DB 71 ATGGTGATCAAGTGTCTCTGTAAACACAGAGATAGACTTTTACTGTGTGGTCTCTTTCC 130
QY 241 AGGGATCATCGAAGTTGGACTATATCCAGGGAATGGCTTCACAGCATCTGGATCA 300
DB 131 AAGGCATCATAAAGATTTGGATTTACATCAAAAGATATGGGCTTTACTGCTATTTGGATT 190
QY 301 CCCCCTTACAGCCAGCTGCCACAGCCACGCATATGGAGATGCTTACCATTGGCTACT 360
DB 191 CTCAGTTGTTGAAACATCTCCGATACACAGCATATGGTTATGCTTATCATGTTGGT 250
QY 361 GGCAGCAGGATATATCTCTCTGAACGAAAACTACGGCATCGCAGATGACTTTGAAGGGC 420
DB 251 GGATGAAGAACATATACAAAAATTAATGAAAACTTTGGTACTGCTGATGATTTGAAGTCTT 310
QY 421 TCTCTCGGCCCTTCTATGAGAGGGGATGTATCTTATGTGCTGATGGTTGCTAACCAT 480
DB 311 TGGCACAAGAAATGGCACGATCGTGATATGTTGTTAATGTGGTATATCGTTTACCAACCAT 370

RESULT 12

US-08-467-831-3

; Sequence 3, Application US/08467831

; Patent No. 5635378

GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-467-831-3

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 8.3e-98;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY	121	CTCGGACTGGCGATCGCAATCCATTATTTCCTCTCTCAGCGATCGATTGCAAGGACGG	180
DB	11	CTGATAAATGGAGATCACAGTCTATTATCAAAATTGCTCAGATGACAGATTGCTAGAACGG	70
QY	181	ATGGGTGCGACGATGGCTTTGTAATCTCGGATCAGAAATCTGTGGTGGAACTGGC	240
DB	71	ATGGTGATACAAGTCTTCCTGTAACACAGATAGACTTTACTGTGGTGGTCTTTCC	130
QY	241	AGGGCATCATCGCAAGTTGGACTATATCCAGGAATGGCTTCACAGCCATCTGGATCA	300
DB	131	AGGGCATCATAAAGAGTTGGATTATCAATCAAGATATGGGCTTTACTGCTATTGGATT	190
QY	301	CCCCGGTTACAGCCAGCTGCCCCAGACCCAGATGAGATGCTACACATGGGTACT	360
DB	191	CTCAGTTGTTGAACAATCCCGATACACAGCATATGGTTATGCTTATCATGGTCTCT	250
QY	361	GGCAGCAGGATATATCTCTCTGAACGAAACTACGSCATCGAGATGACTTGAAGCGC	420
DB	251	GGATGAAGAACATATACAAATTAATGAAACTTTGGTACTGCTGATGATTGAAGTCTT	310
QY	421	TCCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGGTTGCTTAACCAT	480

RESULT 13

US-09-367-891A-5

```
; Sequence 5, Application US/09367891A
; Patent No. 6524816
; GENERAL INFORMATION:
; APPLICANT: RASMUSSEN, PREBEN
; TITLE OF INVENTION: EXPRESSION ELEMENT
; FILE REFERENCE: 078883/0111
; CURRENT APPLICATION NUMBER: US/09/367,891A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: PCT/IB98/00312
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3311
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: amyA gene
US-09-367-891A-5

Query Match      8.2%; Score 143; DB 4; Length 3311;
Best Local Similarity 54.1%; Pred.No. 1.3e-30;
Matches 513; Conservative 0; Mismatches 265; Indels 171; Gaps 4;

QY 482 GGGCTATGATGAGCGGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCCA 541
Db 1141 GGGCTACGGCGGTTCTCATAGCGAGGTGGATTCTCGATCTTCAATCCTCTGAACAGCCA 1200
QY 542 AGACTACTTCCACCGGTTCTGTTTCAATCAAACTATGAAGATCAGACTCAGGTTGAGGA 601
Db 1201 GGATTACTTCCACCGGTTCTGTTTCAATCAAACTATGAAGATCAGACTCAGGTTGAGGA 1260
QY 602 TTGCTGGCTAGGAGATAACACTGTCTCTTCTGCTGATCTCGATACCAACCAAGATGTGT 661
Db 1261 ATGCTGGCTGGCGGATCTCCGACGACATTCGCCGACGTGGACACCAACCAATCCTCAGT 1320
QY 662 CAAGAAATGAATGGTAGCTGGGTGGGATCATTTGGTATCGAACTATCCATT----- 713
Db 1321 TCGGACGTTTTCACGACTGGATCAAGAGCGCTGGTGGGAACTACTCCAGTATGATTGT 1380
QY 714 -----GACGGCTTCGGTAT 727
Db 1381 TCCCGCGGTAAACGCTTTAGGGCTTGCTCTAACTGAAATCGACAGTGCAGTGTCTGGCGT 1440
QY 728 CGACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGTACAAACAGCCGAGCGT 787
Db 1441 CGACACGTTAGCAGCTGGAGAAAGATTTCTGGCCCGGACTTCAACGAAGCTGC---TGC 1497
QY 788 GTACTGTATCGGCGAGGTGCTCGACGGTGTATCGGCTTACACTTGTCCCTACCAAGACT 847
Db 1498 GTGTACCGTCCGCGAGGTGTTCAACGGTGCACCGAGGTACACTGCTGCCCATACCAAGNAGT 1557
QY 848 CATGACGGGCTACTGAATATCC----- 871
Db 1558 GCTGGATGGGCTTCTGAATATCCGATGTGAGTGATTCGGAAGTTCCATCGATCAGGCT 1617
QY 872 -----CATTTACTATCTCACTCTCAACGCTTCAAGTCAACTCCGGCA 915
Db 1618 TTCTGACGATGAGAACAGCTACTCTGCGCTTGATGCAITTCAGTCTGTGGCGGCA 1677
QY 916 GCATGACGACCTCTACAACATGATCAACACCGTCAAATCCGACTGTCCAGACTCAACAC 975
Db 1678 ATCTGGCGGCTTGGCTCAGGCCATCACACCGTGCAGGAGAGCTGCAAGGATTCAATC 1737
QY 976 TCTGGGCACTTGTGTCGAGAACCAAGCAACCCACGGTTTCGTTTC----- 1021
Db 1738 TGCTCGGCAATTTCTTGAATACACGATGCTCGCTTTGCTTGTATGGACACTCTT 1797
QY 1022 -----TTACACCAACGA 1033
Db 1798 TTTGAAGCCCTCATCGATTTGGGATGCTGACACGGACAAACAAACAGGTACCGGATGA 1857
QY 1034 CATAGCCCTCGCAAGAACGTCGACGATTCATTCATCTCAACGACGGAATCCCATCAT 1093
```

```
1858 CCTTGTCTCGCCAAGAATGGTCTCGCTTTCATCATCTCTCGATGGTATTCGGATCAT 1917
1094 CTACCGCGGCCAAGAACAGCACTACGCGCGGGAACGACCCCGGAACCGGGAAGCAAC 1153
1918 CTACACGGGCCAGGAGCAGCACTACGCGGTGATCAGCATCCACAAATCGTGAGGCCGT 1977
1154 CTGGCTCTCGGCTATCCGACCGACGAGAGCTGTACAAGTTAAATTCCTCCGCGCAACGC 1213
1978 CTGGCTGTCTGGTACAATACCGACGCGAGCTGTACCAAGTTTCATCAAGAGGCCAATGG 2037
1214 AATCCGGAACATGATCCATTAGCAAAAGATACAGATTTGTTGACTACAG 1262
2038 CATCCGCAACTGGCTATCAGCCAGAACCCGGAATTCACCTCTCTCCAAG 2086

RESULT 14
US-09-367-891A-6/c
; Sequence 6, Application US/09367891A
; Patent No. 6524816
; GENERAL INFORMATION:
; APPLICANT: RASMUSSEN, PREBEN
; TITLE OF INVENTION: EXPRESSION ELEMENT
; FILE REFERENCE: 078883/0111
; CURRENT APPLICATION NUMBER: US/09/367,891A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: PCT/IB98/00312
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 7432
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-367-891A-6

Query Match      8.2%; Score 143; DB 4; Length 7432;
Best Local Similarity 54.1%; Pred.No. 2.1e-30;
Matches 513; Conservative 0; Mismatches 265; Indels 171; Gaps 4;

QY 482 GGGCTATGATGAGCGGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCCA 541
Db 4777 GGGCTACGGCGGTTCTCATAGCGAGGTGGATTACTCGATCTTCAATCCTCTGAACAGCCA 4718
QY 542 AGACTACTTCCACCGGTTCTGTTTCAATCAAACTATGAAGATCAGACTCAGGTTGAGGA 601
Db 4717 GGATTACTTCCACCGGTTCTGTTTCAATCAAACTATGAAGATCAGACTCAGGTTGAGGA 4658
QY 602 TTGCTGGCTAGGAGATAACACTGTCTCTTGGCTGATCTCGATACCAACCAAGATGTGT 661
Db 4657 ATGCTGGCTGGCGGATCTCCGACGACATTTGCCGAGTGGACACCAACATCTCAGT 4598
QY 662 CAAGAAATGAATGGTAGCTGGGTGGGATCATTTGGTATCGAACTATCCATT----- 713
Db 4597 TCGGACGTTTTCACGACTGATCAAGAGCCTGGTGGGAACTACTCTCAGTATGATTGT 4538
QY 714 -----GACGGCTTCGGTAT 727
Db 4537 TCCCGCGGTAAACGCTTTAGGGCTTGCTCTAACTGAAATCGACAGTGCAGTGGTCTGGCGT 4478
QY 728 CGACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGTACAAACAGCCGAGCGCT 787
Db 4477 CGACACCGTTAAGCAGCTGGAGAAAGATTTCTGGCCCGGACTTCAACGAAGCTGC---TGC 4421
QY 788 GTACTGTATCGGCGAGGTGCTCGACGGTGTATCCGCTTACACTTGTCCCTACCAAGACT 847
Db 4420 GTGTACCGTCCGCGAGGTGTTCAACGGTGCACCGGATACACTGCCCCATACCAAGAGT 4361
QY 848 CATGACGGGCTACTGAATATCC----- 871
```


Search completed: February 12, 2005, 02:00:17
Job time : 331 secs

```
Db 4360 GCTGGATGGCGTTCTGAACATATCCGATGTGAGTGATTCCGAAAGTTCCATCGATCAGGCT 4301
QY 872 -----CATTTACTATCCACTCTCTCAAGCGCTTCAAGTCAACCTCCGGCA 915
Db 4300 TTCTGAGCATGAGAAGACAGCTACTCTCTCGCTTGTATGATTTCAAGTCTGTGCGGGCA 4241
QY 916 GCATGACGACCTCTACAAATGATCAACACCGTCAAAATCCGACTGTCCAGACTCAACAC 975
Db 4240 ATCTCGGGCGTTGGCTCAGGCCATCACACCGTGCAGGAGAGCTGCAAGGATTCATC 4181
QY 976 TCCTGGGACATTCGTGAGAGAACGACACACCCACGGTTTCGCTTC----- 1021
Db 4180 TGCTCGGCAATTTCTTGAGAAATCAGCAATTCGCTTGTGCTTGTATGGACACTCTT 4121
QY 1022 -----TTACACCAACGA 1033
Db 4120 TTTGAAGCCCTCATCGATTTGGGGATGCTGACACGGACAAACACACAGGTACACGGATGA 4061
QY 1034 CATAGCCCTCGCAAGAACGTCGAGCATTCATCATCTCTCAACGACGGAAATCCCATCAT 1093
Db 4060 CCTGTCTCTCGCCAGAAATGGTCTCGCTTTCATCATCTCTCGGATGGTATTCGATCAT 4001
QY 1094 CTACGCGCGCAAGAACAGACACTACGCGCGGGAACGACCCCGGAAACCGCGAAGCAAC 1153
Db 4000 CTACACGGGCGAGGAGCAGCACTACGCGGTGATCAGATCCCAAAATCGTGAGGCCGT 3941
QY 1154 CTGGCTCTCGGGTACCCGACGACAGCGAGCTGTACAGTTAATTTGCCCTCCGCGAACGC 1213
Db 3940 CTGGCTGTCTGGCTACAAATACCGACGCGAGCTGTACAGTTTCAAGAAAGGCCAATGG 3881
QY 1214 AATCCGGAACTATGCCATTAGCAAGATACAGGATTCGTGACCTACAAAG 1262
Db 3880 CATCCGCAACTTGGCTATCAGCCAGAACCCGGAATTCACCTCTCTCCAAG 3832
```

RESULT 15
US-09-189-060B-63
; Sequence 63, Application US/09189060B
; Patent No. 6270968
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Markus
; APPLICANT: Borge, Biderichsen
; TITLE OF INVENTION: Method Of Providing No. 6270968el DNA Sequences
; FILE REFERENCE: 4772.204-US
; CURRENT APPLICATION NUMBER: US/09/189, 060B
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/DK97/00216
; PRIOR FILING DATE: 1997-05-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-189-060B-63

Query Match 4.0%; Score 68.6; DB 3; Length 93;
Best Local Similarity 94.7%; Pred. No. 6e-10;
Matches 71; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 41 GAAGGCATTATGATGGTCGCGTGGTCTCTATTCTGTACGGCTTCAGGTCCGGC 100
Db 4 GATCCCATTTATGATGGTCGCGTGGTCTCTATTCTGTACGGCTTCAGGTCCGGC 63
QY 101 ACCTGCTTTGGTGC 115
Db 64 ACCTGCTTTGGTGC 78

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2005, 21:34:18 ; Search time 949 Seconds

(without alignments)

10781.241 Million cell updates/sec

Title: US-10-820-200-1

Perfect score: 1734

Sequence: 1 tcacatcaagctctcccttc.....aaaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5378673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1734	100.0	1734	18	US-10-820-200-1
2	1707	98.4	1914	15	US-10-653-047-5355
3	1011.2	58.3	3285	15	US-10-228-063-46
4	979	56.5	2443	18	US-10-815-495-17
5	908.2	52.4	2520	18	US-10-815-495-21
6	806	46.5	1389	18	US-10-653-047-3775
7	786	45.3	1488	14	US-10-213-990-41
8	648.2	37.4	1893	14	US-10-213-990-44
9	439	25.3	1389	18	US-10-653-047-3775
10	374.2	21.6	483	18	US-10-653-047-3837
11	369.2	21.3	4190	15	US-10-128-590-28
12	369.2	21.3	4190	16	US-10-128-597A-28
13	369.2	21.3	4190	17	US-10-128-578B-28
14	368.8	21.3	605	18	US-10-653-047-3796
15	344.2	19.9	1048	18	US-10-653-047-3776
16	317.8	18.3	1644	14	US-10-213-990-47
17	308	17.8	1926	14	US-10-213-990-40
18	266.2	15.4	2250	14	US-10-213-990-43
19	252.4	14.6	1502	17	US-10-369-493-36156
20	244.8	14.1	1707	14	US-10-213-990-46
21	236.4	13.6	619	18	US-10-653-047-5061
22	220.6	12.7	292	18	US-10-653-047-3878
23	206.6	11.9	1602	18	US-10-653-047-187
24	196.2	11.3	1539	17	US-10-369-493-25910
25	174.4	10.1	1892	17	US-10-369-493-46366
26	165.4	9.5	592	18	US-10-653-047-3791
27	138.6	8.0	855	18	US-10-425-115-106984
28	136.6	7.9	1479	17	US-10-369-493-46410
29	129.2	7.5	173	18	US-10-653-047-3995
30	77.4	4.5	1509	15	US-10-081-872-107
31	77.4	4.5	1509	17	US-10-385-305-107
32	70.6	4.1	839	18	US-10-653-047-3800
33	54.2	3.1	1452	17	US-10-369-493-24284
34	53.2	3.1	2731748	18	US-10-297-465A-1
35	53.2	3.1	2731748	18	US-10-297-465A-1
36	52.4	3.0	77536	10	US-09-940-316B-1
37	51.8	3.0	2364	18	US-10-473-687-3
38	51.4	3.0	985	18	US-10-363-345A-31407
39	51.4	3.0	985	18	US-10-363-345A-31408
40	51.2	3.0	520	14	US-10-184-644-332
41	51.2	3.0	520	14	US-10-184-634-332
42	51.2	3.0	867	9	US-09-216-393-340
43	51.2	3.0	867	9	US-09-216-393-342
44	51.2	3.0	867	16	US-10-321-856-340
45	51.2	3.0	867	16	US-10-321-856-342

Query Match 100.0%; Score 1734; DB 18; Length 1734;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

RESULT 1

US-10-820-200-1
; Sequence 1, Application US/10820200
; Publication No. US20040229764A1
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Pedersen, Sven
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Fungamyl-like Alpha-Amylase Variants
; FILE REFERENCE: 5835.200-US
; CURRENT APPLICATION NUMBER: US/10/820,200
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US/09/710,339
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/165,786
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Aspergillus Oryzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(1547)
; OTHER INFORMATION: mat_peptide
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (114)..(1733)
US-10-820-200-1

Sequence 28, Appl
Sequence 28, Appl
Sequence 3796, Ap
Sequence 3776, Ap
Sequence 47, Appl
Sequence 40, Appl
Sequence 43, Appl
Sequence 36156, A
Sequence 46, Appl
Sequence 5061, Ap
Sequence 3878, Ap
Sequence 187, App
Sequence 25910, A
Sequence 46366, A
Sequence 3791, Ap
Sequence 106984,
Sequence 46410, A
Sequence 3995, Ap
Sequence 107, App
Sequence 107, App
Sequence 3800, Ap
Sequence 24284, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 31407, A
Sequence 31408, A
Sequence 332, App
Sequence 332, App
Sequence 340, App
Sequence 340, App
Sequence 342, App

```
QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAAATAAACCCCAAGAGGCATTTATGATGGTGG 60
Db |||
1 TCACATCAAGCTCTCCCTTCTCTGAACAAATAAACCCCAAGAGGCATTTATGATGGTGG 60
QY 61 CGTGGTGGTCTCTATTTCTGTATCGGCTTTCAGGTGCGGCACTGCTGGCTGCAAGCG 120
Db |||
61 CGTGGTGGTCTCTATTTCTGTATCGGCTTTCAGGTGCGGCACTGCTGGCTGCAAGCG 120
QY 121 CTGCGGACTGCGGATCGCAATCCATTTATTTCTTCTACGGATCGATTTGCAAGGCGG 180
Db |||
121 CTGCGGACTGCGGATCGCAATCCATTTATTTCTTCTTCTACGGATCGATTTGCAAGGCGG 180
QY 181 ATGGGTTCGACGACTGCGACTTGTATATCTGCGGATCAGAAATACCTGTGTGGTGAACATGGC 240
Db |||
181 ATGGGTTCGACGACTGCGACTTGTATATCTGCGGATCAGAAATACCTGTGTGGTGAACATGGC 240
QY 241 AGGGCATCATCGCAAGTTGGACTATATCCAGGGAATGGGCTTTCAGGCCATCTGGATCA 300
Db |||
241 AGGGCATCATCGCAAGTTGGACTATATCCAGGGAATGGGCTTTCAGGCCATCTGGATCA 300
QY 301 CCCCGTTACAGCCAGCTGCCCGACCAACCGCATATGGAGATGCCTACCATGGCTACT 360
Db |||
301 CCCCGTTACAGCCAGCTGCCCGACCAACCGCATATGGAGATGCCTACCATGGCTACT 360
QY 361 GGCAGCAGGATATATATCTCTGAACCAAAACTACGGCACTGCAGATGACTTTGAAGGCGC 420
Db |||
361 GGCAGCAGGATATATATCTCTGAACCAAAACTACGGCACTGCAGATGACTTTGAAGGCGC 420
QY 421 TCTCTTGGCCCTTATGAGAGGGGATGATATCTTATGGTTCGATGCGTTGTAACCATTA 480
Db |||
421 TCTCTTGGCCCTTATGAGAGGGGATGATATCTTATGGTTCGATGCGTTGTAACCATTA 480
QY 481 TGGGCTATATGAGGCGGTAGCTCAGTCGATACAGTGTGTTTAAACCGTTTCAGTTCCC 540
Db |||
481 TGGGCTATATGAGGCGGTAGCTCAGTCGATACAGTGTGTTTAAACCGTTTCAGTTCCC 540
QY 541 AAGACTACTTCCACCGGTTCTGTTTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
Db |||
541 AAGACTACTTCCACCGGTTCTGTTTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
QY 601 ATTGCTGGCTAGGAGATAACATGTCTCTTCCCTGATCTCGATACCAACCAAGGATGG 660
Db |||
601 ATTGCTGGCTAGGAGATAACATGTCTCTTCCCTGATCTCGATACCAACCAAGGATGG 660
QY 661 TCAAGATCAATGGTACGACTGGTGGGATCATTTGGTATCGAATCTACTCCATTTGAGCGCC 720
Db |||
661 TCAAGATCAATGGTACGACTGGTGGGATCATTTGGTATCGAATCTACTCCATTTGAGCGCC 720
QY 721 TCCGTATCGACACAGTAAACACAGTCCAGAGGACTTCTGGCCCGGGTACAAACAAGCGG 780
Db |||
721 TCCGTATCGACACAGTAAACACAGTCCAGAGGACTTCTGGCCCGGGTACAAACAAGCGG 780
QY 781 CAGGCGTGTACTGTATTCGCGAGGTTGCTCGACGGTGATCCGGCCTACACTTTCCTTACC 840
Db |||
781 CAGGCGTGTACTGTATTCGCGAGGTTGCTCGACGGTGATCCGGCCTACACTTTCCTTACC 840
QY 841 AGAAGCTATGACGCGGTACTGAACTATCCCATTTTACTATCCACTCTTCAACGCGCTTCA 900
Db |||
841 AGAAGCTATGACGCGGTACTGAACTATCCCATTTTACTATCCACTCTTCAACGCGCTTCA 900
QY 901 AGTCAACCTCCGCGACGATGGACGACTCTTACAACATGATCAACACCGTCAAAATCCGACT 960
Db |||
901 AGTCAACCTCCGCGACGATGGACGACTCTTACAACATGATCAACACCGTCAAAATCCGACT 960
QY 961 GTCCAGACTCAACACTCTCTGGGACATTCGTGAGAACCAACGACACCGGTTGCTTT 1020
Db |||
961 GTCCAGACTCAACACTCTCTGGGACATTCGTGAGAACCAACGACACCGGTTGCTTT 1020
QY 1021 CTTACACCAACGACATAGCCCTCGCCAAAGACGTCGACGATTCATCTCCTCAAGCAGG 1080
Db |||
1021 CTTACACCAACGACATAGCCCTCGCCAAAGACGTCGACGATTCATCTCCTCAAGCAGG 1080
```

```
QY 1081 GAATCCCCCATCATCTACCCCGGCCAAGAACAGACTACGCGCGCGGAAACGACCCGCGA 1140
Db |||
1081 GAATCCCCCATCATCTACCCCGGCCAAGAACAGACTACGCGCGCGGAAACGACCCGCGA 1140
QY 1141 ACCGGAAGCAACCTGGCTCTCGGGCTACCCGACGACAGGAGCTGTACAAGTTAAATTG 1200
Db |||
1141 ACCGGAAGCAACCTGGCTCTCGGGCTACCCGACGACGAGCTGTACAAGTTAAATTG 1200
QY 1201 CTTCCGCGAAACGCAATCCGAACTATGCAATTTAGCAAAAGATACAGGATTTGTAAGCTACA 1260
Db |||
1201 CTTCCGCGAAACGCAATCCGAACTATGCAATTTAGCAAAAGATACAGGATTTGTAAGCTACA 1260
QY 1261 AGAATCGGCCCATCTTCAAAAGACGACACAGCATCGCCATGCGCAAGGCGACAGATGGT 1320
Db |||
1261 AGAATCGGCCCATCTTCAAAAGACGACACAGCATCGCCATGCGCAAGGCGACAGATGGT 1320
QY 1321 CGCAGATCGTACTATCTTGTCCAAAGGGTGCCTTCGGGTGATTCGTATACCCCTCTCCT 1380
Db |||
1321 CGCAGATCGTACTATCTTGTCCAAAGGGTGCCTTCGGGTGATTCGTATACCCCTCTCCT 1380
QY 1381 TGAGTGGTGGGTTTACACAGCCGCGCAGCAATTCACGAGGTCAATTTGGCTGCACGACCG 1440
Db |||
1381 TGAGTGGTGGGTTTACACAGCCGCGCAGCAATTCACGAGGTCAATTTGGCTGCACGACCG 1440
QY 1441 TGACGGTTGGTTCGGATGGAATGTGCTTCTATGCGCAGGTGGGCTACCTAGGGTAT 1500
Db |||
1441 TGACGGTTGGTTCGGATGGAATGTGCTTCTATGCGCAGGTGGGCTACCTAGGGTAT 1500
QY 1501 TGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGTTGAGA 1560
Db |||
1501 TGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGTTGAGA 1560
QY 1561 GTATATGATGGTACTGCTATTCAATCTGCGATGACAGTGTGAGTTTGTAGTTGATGTACA 1620
Db |||
1561 GTATATGATGGTACTGCTATTCAATCTGCGATGACAGTGTGAGTTTGTAGTTGATGTACA 1620
QY 1621 GTTGGAGTCTGTACTGCTATCCCTTATCTCTCGATTTGTTTTCGAACCCCTAATG 1680
Db |||
1621 GTTGGAGTCTGTACTGCTATCCCTTATCTCTCGATTTGTTTTCGAACCCCTAATG 1680
QY 1681 CCAAGCAGCTAGTCTATTATAGGAAAAAAGGAAAAAAAAAAAAAAAAAAAAA 1734
Db |||
1681 CCAAGCAGCTAGTCTATTATAGGAAAAAAGGAAAAAAAAAAAAAAAAAAAAA 1734
```

RESULT 2

```
; US-10-653-047-5355
; Sequence 5355, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5355
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
; US-10-653-047-5355
```

Query Match 98.4%; Score 1707; DB 18; Length 1914; ;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACATCAAGCTCTCCCTCTCTGAAACAATAAACCACAGAGGATTTATGATGGTCG 60
DB 3 TCACATCAAGCTCTCCCTCTCTGAAACAATAAACCACAGAGGATTTATGATGGTCG 62
QY 61 CGTGGTGGTCTATTTCTGTATCGGCTTCAGTCCGGCACTGCTTTGGCTGCAACGC 120
DB 63 CGTGGTGGTCTATTTCTGTATCGGCTTCAGTCCGGCACTGCTTTGGCTGCAACGC 122
QY 121 CTGGGACTGGCGATCGCAATCAATTTATTTCTCTCAGGATCGATTTGCAAGACGG 180
DB 123 CTGGGACTGGCGATCGCAATCAATTTATTTCTCTCAGGATCGATTTGCAAGACGG 182
QY 181 ATGGGTCGACGACTGGGACTTTGTAATCTCGGATCAGAAATACTGTGGTGGAAATGGC 240
DB 183 ATGGGTCGACGACTGGGACTTTGTAATCTCGGATCAGAAATACTGTGGTGGAAATGGC 242
QY 241 AGGGCATCATCGAAAGTTGGACTATATCCAGGAAATGGGCTTTCAGGCCATCTGGATCA 300
DB 243 AGGGCATCATCGAAAGTTGGACTATATCCAGGAAATGGGCTTTCAGGCCATCTGGATCA 302
QY 301 CCCCCGTTACAGCCAGCTGCCCCAGACACCCGATATGAGATGCTTACCATGGCTACT 360
DB 303 CCCCCGTTACAGCCAGCTGCCCCAGACACCCGATATGAGATGCTTACCATGGCTACT 362
QY 361 GGACAGAGGATATATCTCTGAAACGAACTACGGCACTGAGATGACTTCAAGGCGC 420
DB 363 GGACAGAGGATATATCTCTGAAACGAACTACGGCACTGAGATGACTTCAAGGCGC 422
QY 421 TCTCTTTCGGCCCTTCATGAGAGGGGATGATCTTATGGTTCGATGGTTCGCTAACCAT 480
DB 423 TCTCTTTCGGCCCTTCATGAGAGGGGATGATCTTATGGTTCGATGGTTCGCTAACCAT 482
QY 481 TGGGCTATGATGAGCGGGTAGCTCAGTCGATTCAGTGTGTTAAACCGTTCAGTCCC 540
DB 483 TGGGCTATGATGAGCGGGTAGCTCAGTCGATTCAGTGTGTTAAACCGTTCAGTCCC 542
QY 541 AAGACTACTTCCACCCGTTCTGTTTCATTCAAACTATGAAGATCAGACTCAGGTTGAGG 600
DB 543 AAGACTACTTCCACCCGTTCTGTTTCATTCAAACTATGAAGATCAGACTCAGGTTGAGG 602
QY 601 ATTGCTGGCTAGGATTAACATCTCTCTTCCTGCTGATCTCGATACCAACCAAGGATG 660
DB 603 ATTGCTGGCTAGGATTAACATCTCTCTTCCTGCTGATCTCGATACCAACCAAGGATG 662
QY 661 TCAAGAAATGAATGATGATGAGGATGATGATGATGATGATGATGATGATGATGATG 720
DB 663 TCAAGAAATGAATGATGATGAGGATGATGATGATGATGATGATGATGATGATGATG 722
QY 721 TCCGATTCGACACAGTAAACAGCTCCAGAGGACTTCTGGCCCGGGTACAAAGACCG 780
DB 723 TCCGATTCGACACAGTAAACAGCTCCAGAGGACTTCTGGCCCGGGTACAAAGACCG 782
QY 781 CAGGCGGTACTGATCTCGGCGAGGTGCTCGACGGTATCGGCGCTACACTTGTCCCTACC 840
DB 783 CAGGCGGTACTGATCTCGGCGAGGTGCTCGACGGTATCGGCGCTACACTTGTCCCTACC 842
QY 841 AGAACGTATGAGACGGGTAAGTAACTATCCCAATTTACTATCCACTCTCCCAAGCCCTTCA 900
DB 843 AGAACGTATGAGACGGGTAAGTAACTATCCCAATTTACTATCCACTCTCCCAAGCCCTTCA 902
QY 901 AGTCAACCTCCGGCAGCATGGAGCTCTTACAACTATGATCAACCGTCAAAATCCGACT 960
DB 903 AGTCAACCTCCGGCAGCATGGAGCTCTTACAACTATGATCAACCGTCAAAATCCGACT 962
QY 961 GTCCAGACTCAACACTCTCTGGGACATTCGTTCGAGAACCAACCAACCCAGGTTGGCTT 1020
DB 963 GTCCAGACTCAACACTCTCTGGGACATTCGTTCGAGAACCAACCAACCCAGGTTGGCTT 1022
QY 1021 CTTTACACCAACGATAGCCCTTCGCCAAGAACGTCGAGCATTCATCATCTCTCAACGACG 1080

DB 1023 CTTTACACCAACGACATAGCCCTCGCCAAAGACGTGCGACATTCATCATCTCTCAACGACG 1082
QY 1081 GAATCCCATCATCTACCGCGGCCAAGAACAGCACTACCCCGCGGAAACGACCCGCGGA 1140
DB 1083 GAATCCCATCATCTACCGCGGCCAAGAACAGCACTACCCCGCGGAAACGACCCGCGGA 1142
QY 1141 ACCGGAAGCAACCTGGCTCTCGGGTACCCGACCGACAGGAGCTGTCAAGTTAATTG 1200
DB 1143 ACCGGAAGCAACCTGGCTCTCGGGTACCCGACCGACAGGAGCTGTCAAGTTAATTG 1202
QY 1201 CTTCCGCGAAGCAATCCGGAATAGCCATTTAGCAAAAGATACAGGATTCGTGACCTACA 1260
DB 1203 CTTCCGCGAAGCAATCCGGAATAGCCATTTAGCAAAAGATACAGGATTCGTGACCTACA 1262
QY 1261 AGAATCTGGCCCATCTACAAAGACGACACAACGATCGCCATGCGCAAGGCGACAGATGGGT 1320
DB 1263 AGAATCTGGCCCATCTACAAAGACGACACAACGATCGCCATGCGCAAGGCGACAGATGGGT 1322
QY 1321 CCGAGATCGTACTATCTTGTCCAAAGGGTCTTTCGGGTGATTCGTATACCTCTCCT 1380
DB 1323 CCGAGATCGTACTATCTTGTCCAAAGGGTCTTTCGGGTGATTCGTATACCTCTCCT 1382
QY 1381 TGAGTGGTCCGGTTCACAGCGCGCCAGCAATTTGACGAGGTCATTTGGCTGCACGACCG 1440
DB 1383 TGAGTGGTCCGGTTCACAGCGCGCCAGCAATTTGACGAGGTCATTTGGCTGCACGACCG 1442
QY 1441 TGACGGTTCGGTTCGGATGGAATGTCCCTTATGTCAGGTCGGCTACCTAGGGTAT 1500
DB 1443 TGACGGTTCGGTTCGGATGGAATGTCCCTTATGTCAGGTCGGCTACCTAGGGTAT 1502
QY 1501 TGTATCCGACTCAGAAAGTTGGCAGGTAGCAAGATCTGTAGTACTCTGTAAGGGTGGAGA 1560
DB 1503 TGTATCCGACTCAGAAAGTTGGCAGGTAGCAAGATCTGTAGTACTCTGTAAGGGTGGAGA 1562
QY 1561 GTATATGATGGTACTGCTATTCATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCG 1620
DB 1563 GTATATGATGGTACTGCTATTCATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCG 1622
QY 1621 GTTGGAGTCTGTTACTGCTCTATCCCTTATATCTCTTCCATTTGTTTTCGAAACCTTAATG 1680
DB 1623 GTTGGAGTCTGTTACTGCTCTATCCCTTATATCTCTTCCATTTGTTTTCGAAACCTTAATG 1682
QY 1681 CCAAGCAGCTAGTCTATTTATAGGAAA 1707
DB 1683 CCAAGCAGCTAGTCTATTTATAGGAAA 1709

RESULT 3

US-10-228-063-46
; Sequence 46, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3285
; TYPE: DNA
; ORGANISM: Aspergillus shiroyami
US-10-228-063-46

Query Match 58.3%; Score 1011.2; DB 15; Length 3285;
Best Local Similarity 81.9%; Pred. No. 2.6e-302;
Matches 1166; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

QY 114 GCAAGCCTCGGAGCTGGCGATCGGAATCCATTTATTTCTCTCAGCGATTCGATTCGA 173
DB 1 GGCACCCCGGCGACTGGCGCTCCAGTCCATCTACTTCTCTCTCACCAGCGCTTCGCC 60

QY 174 AGGACGGGATGGGTTCAGACGCTGCGACTTGTAAATCTGCGGATCAGAAAATCTGTGTGGA 233
DB |||||
QY 61 CGCACGACGGCTCCACACCGCCACCTGCACACCGCCGACAGAGTACTTGGGGGC 120
DB |||||
QY 234 ACATGCGAGGGGATCATCGACAAGTTGGATATATCCAGGGAAATGGGCTTCAAGCATTC 293
DB |||||
QY 121 ACCTGCGAGGGGATCATCGACAAGCTCGACTATCATCCAGGGATGGGCTTCAAGCATTC 180
DB |||||
QY 294 TGGATCACCCCGTTTACAGCCGAGCTGCCCGAGACACCGCATATGGAGATCGCTACCAT 353
DB |||||
QY 181 TGGATCACCCCGTTGACCGCCGAGCTCCCGGAGACACCGCATAGCGGACGCTTACCAT 240
DB |||||
QY 354 GGCTACTGCGAGCAGGATATATCTCTCTGAAGAAATACGGGACTCGAGTACTGACTTG 413
DB |||||
QY 241 GGTACTGCGAGCAGGATATCTACTCTCTCAAGGAACTAGCGGACCGCGGAGCATTC 300
DB |||||
QY 414 AAGGCGCTCTCTTTCGCGCTTCATAGAGGGGATGATCTTATGGTGCATGTGTTGCT 473
DB |||||
QY 301 AAGGCGCTCTCTTTCGCGCTTCACAGAGCGGATGTACTCTCATGTGGACGTGGTGGCC 360
DB |||||
QY 474 AACCATATGGGTATGATGGAGGGGTAGCTCAGTCTGATTACAGTGTGTTAAACGTTTC 533
DB |||||
QY 361 AACCATATGGGTATGATGGAGGGGTAGCTCCTCGTGGACTCTCGGTGTTCAAGCGGTTTC 420
DB |||||
QY 534 AGTTCCCAAGACTTCTCCACCGGTTCTGTTTCAATTCAAAATATGAAAGATCAGACTCAG 593
DB |||||
QY 421 TCTCTCCAGGACTTACTTCCACCGGTTCTGCTTCAATCCAGAACTACGAGGACAGACCGAG 480
DB |||||
QY 594 GTTGAGGATTTGGGTAGGATTAACACTGTCTCTTCCCTGCTGATCTCGATACCAACCAAG 653
DB |||||
QY 481 GTTGAGGACTGTGCTCGGCGACAAACCGGTGCTCTCCCGGACTCGACACCAACCAAG 540
DB |||||
QY 654 GATGGGTCAAGATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCAT 713
DB |||||
QY 541 GACGTGGTGAAGAACGAGTGGTACGACTGGGTGGGCTCCCTCGTCTCAACTACTCCATC 600
DB |||||
QY 714 GACGGCTCCGTATCGACACAGTAAACACGCTCCAGAGGACTTCTGGCGCGGGTACAAAC 773
DB |||||
QY 601 GACGGCTCCGATCGACACCGTGAAGCAGCTGCAGAGGACTTCTGGCGGGGCTACAAAC 660
DB |||||
QY 774 AAAGCGCAGGGGTGTAATGATTCGGGAGGTGCTCGAGGTGATTCGGGCTTACACTTGT 833
DB |||||
QY 661 AAAGCGCGCGGTGTAATGATTCGGGAGGTGCTCGAGGTGATTCGGGCTTACACTTGT 720
DB |||||
QY 834 CCTTACCAAGCTGATGACGCGGTACTGAAGTATCCCATTTACTATCCATCTCTCAAC 893
DB |||||
QY 721 CCGTACCAAGCTGATGACGCGGTGCTGAAGTATCCCATTTACTATCCATCTCTCAAC 780
DB |||||
QY 894 GCCTTCAAGTCAACCTCCGGCAGCATGGACGACCTCTCAACATGATCAACACCGTCAAA 953
DB |||||
QY 781 GCCTTCAAGTCAACCTCCGGCAGCATGGACGACCTCTCAACATGATCAACACCGTGAAG 840
DB |||||
QY 954 TCCGACTGTCAGACTCAACACTCTTGGGCAATTTGTCGAGAACCAACCAACCAACGG 1013
DB |||||
QY 841 TCCGACTGTCAGACTCAACACTCTTGGGCAATTTGTCGAGAACCAACCAACCAACGG 900
DB |||||
QY 1014 TTGCGCTTCTTACCAACGACATAGCCCTCGCCAGAAAGCTGCGAGCATTCATCATCTTC 1073
DB |||||
QY 901 TTGCGCTTCTTACCAACGACATAGCCCTCGCCAGAAAGCTGCGAGCATTCATCATCTTC 960
DB |||||
QY 1074 AACGCGGAATCCCATCATCTTACGCGCGCCAGAACAGCATAGCGCGCGGGAACGAC 1133
DB |||||
QY 961 AACGCGGATCCCGATCATCTTACGCGCGCCAGGAGCAGCATAGCGCGCGGGAACGAC 1020
DB |||||
QY 1134 CCGGGAACCGGGAAGCAACTTGGGCTCTCGGGCTACCGGACGAGCAGGCTGTACAG 1193
DB |||||
QY 1021 CCGGCAACCGGAGGCACTTGGGCTCTCGGGCTACCGGACGAGCTCCGAGCTGTACAG 1080
DB |||||
QY 1194 TTAATTCCTTCGCGAACGCAATCCGGAATATGCCATTTAGCAAGATACAGATTCGTG 1253
DB |||||
QY 1081 CTATCGCTTCGCGCAACGCCATTCGCAACTAGCCATCTCCAGAGACACCGGCTTCGTG 1140
DB |||||

QY 1254 ACCTACAAGAACTGCCCATCTTACAAGACGACACCAACGATCGCCATCGCAAGGACACA 1313
DB |||||
QY 1141 ACCTACAAGAACTGCCCATCTTACAAGACGACACCAACGATCGCCATCGCAAGGACACC 1200
DB |||||
QY 1314 GATGGGTGCGAGATCGTGAATCTTGTTCACAAAGGGTGTCTTCGGGTGATTCGTATACC 1373
DB |||||
QY 1201 GACGGCTCCAGATCGTGAATCTTCTTCAACAAAGGGCGCTTCGGGAGCTCTTACACC 1260
DB |||||
QY 1374 CTCTCTTGTAGTGGTGGGTTTACACGCGCGGAGCAATTCACGAGGTCAATTTGGCTGC 1433
DB |||||
QY 1261 CTCTCTCTTCTCGGCGCGGCTTACACCGCGCGGAGAGCTCACGAGGTGATTCGGTGC 1320
DB |||||
QY 1434 ACGACCGTGAAGTGGTTCGGATGGAATGTGCTTCTTATGCGAGGTGGGCTACCT 1493
DB |||||
QY 1321 ACCACCGTGAAGTGGTTCGGAGCGCAACGTTCGGGTGCGATGCGCGGCGCTCCCG 1380
DB |||||
QY 1494 AGGTATTTGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTG 1537
DB |||||
QY 1381 CGCGTCTTACCCGACGAGAAAGCTCGCCGCTCCAAAGATATG 1424
DB |||||

RESULT 4
US-10-815-495-17
; Sequence 17, Application US/10815495
; Publication No. US20040191864A1
; GENERAL INFORMATION:
; APPLICANT: Novozymes Biotech, Inc.
; APPLICANT: Connelly, Mariah
; APPLICANT: Brody, Howard
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient
; TITLE OF INVENTION: Mutants Of Aspergillus Niger
; FILE REFERENCE: 10345-200-US
; CURRENT APPLICATION NUMBER: US/10/815,495
; CURRENT FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: n=a,c,g or t
US-10-815-495-17

Query Match 56.5%; Score 979; DB 18; Length 2443;
Best Local Similarity 74.4%; Pred. No. 2.2e-292;
Matches 1609; Conservative 0; Mismatches 10; Indels 544; Gaps 8;

QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTGC 60
DB |||||
QY 151 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGGCAATTTATGATGGTGC 210
DB |||||
QY 61 CGTGTGTGCTCTATTTCTGTACGGCTTCAGGTGCGCGCACCTGCTTTGGTGCACACGC 120
DB |||||
QY 211 CGTGTGTGCTCTATTTCTGTACGGCTTCAGGTGCGCGCACCTGCTTTGGTGCACACGC 270
DB |||||
QY 121 CTGCGGACTGCGGATCGCAATTCATTTATTTCTTCTCACGGATCGATTTGCAAGACGG 180
DB |||||
QY 271 CTGCGGACTGCGGATCGCAATTCATTTATTTCTTCTCACGGATCGATTTGCAAGACGG 330
DB |||||
QY 181 ATGGGTCCAGCAGCTCGGACTTGTAACTACCTGCGGATCAGGTGTTGTTACCTACTAGCTT 216
DB |||||
QY 331 ATGGGTCCAGCAGCTCGGACTTGTAACTACCTGCGGATCAGGTGTTGTTACCTACTAGCTT 390
DB |||||
QY 217 -----AGAAATATCTGTGTGGAACATGGCAGGGC 245
DB |||||
QY 391 TCAGAAAGAGGAACTGAATGACTTGTATATAGAAATATCTGTGTGGAACATGGCAGGGC 450
DB |||||
QY 246 ATCATCGACA----- 255
DB |||||
QY 451 ATCATCGACAAGGTAATTTGCCCTTTTATCAAAAAAGAGAAAGAGAGAGAGAGAGAGAG 510
DB |||||

QY 256 -----AGTTGGACTATATCCAGGGAATGGG 280
Db TAAATAAAGAACTCTAGTCTTAACCATACATAGTTGGACTATATCCAGGGAATGGG 570
QY 281 CTTACAGCCATCTGGATCACCCTCGTTACAGCCAGCTGCCCGACAGCACCGCATATGG 340
Db CTTACAGCCATCTGGATCACCCTCGTTACAGCCAGCTGCCCGACAGCACCGCATATGG 630
QY 341 AGATGCTTACCATGGCTACTGGCAGCAGGATAT- 373
Db AGATGCTTACCATGGCTACTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC 690
QY 374 -----ATACTCTTGAACGAAA 391
Db TGTCATCTTTTACATCAATATGAACCTTAAGTTAGTATCTCTCTGAACGAAA 750
QY 392 CTACGACATCGAGATGACTTGAAGCGCTCTCTCGGCCCTTCATGAGAGGGGATGTA 451
Db CTACGACATCGAGATGACTTGAAGCGCTCTCTCGGCCCTTCATGAGAGGGGATGTA 810
QY 452 TCTTATGGTGGATGTTGGTCTTAACCATAT- 481
Db TCTTATGGTGGATGTTGGTCTTAACCATATGTTCTGGTCTCTTGCACCTGACTTCGCG 870
QY 482 -----GGGCTATGATGGAGCGGTAGC 503
Db GATATGGTTCATTTTCACTAGTCAATAGATATATCAGGGCTATGATGGAGCGGTAGC 930
QY 504 TCAGTCGATACAGTGTGTTAAACCGTTCAGTTCCTCAAGACTACTTCCACCGGTTCTGT 563
Db TCAGTCGATACAGTGTGTTAAACCGTTCAGTTCCTCAAGACTACTTCCACCGGTTCTGT 990
QY 564 TTCAATTTCAAACTATCAAGATCAGACTCAGGTTGAGGATTTGCTGGCTAGGAGATAACACT 623
Db TTCAATTTCAAACTATCAAGATCAGACTCAGGTTGAGGATTTGCTGGCTAGGAGATAACACT 1050
QY 624 GTCTCTTGGCTGATCTCGATACCAACAGGATGTGTTCAAGATGAATGGTACGACTGG 683
Db GTCTCTTGGCTGATCTCGATACCAACAGGATGTGTTCAAGATGAATGGTACGACTGG 1110
QY 684 GTGGGATCATTTGGTATCGAATCTCA- 711
Db GTGGGATCATTTGGTATCGAATCTCAAGATATTTCTCCCTCATTTCAACATGG 1170
QY 712 -----TTGACGGCTCTCCGTATCGACACAGTAAACACGT 745
Db CTGATCGATGATCTTACGAAATCAGTTGACGGCTCTCCGTATCGACACAGTAAACACGT 1230
QY 746 CCAGAGGACTTCTGGCCCGGGTACAAACAGCCGAGCGGTGTAAGTATCGCGAGGT 805
Db CCAGAGGACTTCTGGCCCGGGTACAAACAGCCGAGCGGTGTAAGTATCGCGAGGT 1290
QY 806 GCTCGAGCGGTATCCGGCTTACACTTGTCCCTACCAAGAGCTCATCGAGCGGTACTGAA 865
Db GCTCGAGCGGTATCCGGCTTACACTTGTCCCTACCAAGAGCTCATCGAGCGGTACTGAA 1350
QY 866 CTATCCCAT- 874
Db CTATCCCATGATGGTTCTTCCCAACCATGAGCCTTCTTCAAGTCTCATCTCTCAACGAA 1410
QY 875 -----TTACTATCACTCTTCAACGCTTCAAGTCAACCTCCGGCAGCATG 920
Db ACCGGTAAACACCATGTTACTATCACTCTTCAACGCTTCAAGTCAACCTCCGGCAGCATG 1470
QY 921 GAGGACTCTAACAATGATCAACAGCTCAATTCGACTGTCCAGACTCAACACTCTCTG 980
Db GAGGACTCTAACAATGATCAACAGCTCAATTCGACTGTCCAGACTCAACACTCTCTG 1530
QY 981 GGCACATTTGTCAGAACACGACCAACCCACGGTTCGCTTC- 1021
Db GGCACATTTGTCAGAACACGACCAACCCACGGTTCGCTTCGTAAGTCTTCCCTTTTATT 1590

QY 1022 -----TTACACCAACGACA 1035
Db TTCCGTTCCAAATTTCCACAGAACCCACCTAAACAGAGCAAAAGTTTACACCAACGACA 1650
QY 1036 TAGCCCTCGCCCAAGAACGTCGAGCATTCATCTCTCAACGACGGAATCCCCATCATCT 1095
Db TAGCCCTCGCCCAAGAACGTCGAGCATTCATCTCTCAACGACGGAATCCCCATCATCT 1710
QY 1096 AGCCGGCCCAAGAACAGCACTACGCGCGGCGGAAAACGACCCCGGAACCCGGAACCACT 1155
Db AGCCGGCCCAAGAACAGCACTACGCGCGGCGGAAAACGACCCCGGAACCCGGAACCACT 1770
QY 1156 GGGCTCTCGGGCTACCGGACCGAGCGAGCTGTACAGTTAATTTGCTCGCGGAACGCAA 1215
Db GGGCTCTCGGGCTACCGGACCGAGCGAGCTGTACAGTTAATTTGCTCGCGGAACGCAA 1830
QY 1216 TCCGGAATCTATGCCATTTAGCAAGATACAGGATTCGTGACCTACA- 1260
Db TCCGGAATCTATGCCATTTAGCAAGATACAGGATTCGTGACCTACAAGGTTAAGCAACACT 1890
QY 1261 ----- 1260
Db CTAAGCATACCTTAATGGCTATCTTACAGTATCTGACACAAGAGACTAATCACTGGCA 1950
QY 1261 -----AGAACTGCGCCATCTACAAAGACGACACACGATCGCCATCGCAAGGGCACAGAT 1316
Db ATACAGAACTGCGCCATCTACAAAGACGACACACGATCGCCATCGCAAGGGCACAGAT 2010
QY 1317 GGGTCGACATCGTGAATCTTTGTCACAAAGGGTCTTCGGGTGATTCGTATACCTC 1376
Db GGGTCGACATCGTGAATCTTTGTCACAAAGGGTCTTCGGGTGATTCGTATACCTC 2070
QY 1377 TCTTTAGTGTGCGGGTTTACACGCGCCGACGCAATTCACGAGGTCATTTGGCTGCACG 1436
Db TCTTTAGTGTGCGGGTTTACACGCGCCGACGCAATTCACGAGGTCATTTGGCTGCACG 2130
QY 1437 ACCGTGACGGTGGTTCGAGTGAATGTGCTTCCTATGCGCAGTGGGCTACCTAGG 1496
Db ACCGTGACGGTGGTTCGAGTGAATGTGCTTCCTATGCGCAGTGGGCTACCTAGG 2190
QY 1497 GTATTGTATCCCACTGAGAAAGTTGCGAGTACAGATCTGTAGTAGCTCGTGAAGGGTG 1556
Db GTATTGTATCCCACTGAGAAAGTTGCGAGTACAGATCTGTAGTAGCTCGTGAAGGGTG 2250
QY 1557 GAGGTATATGATGGTACTGCTATTCAATCTGGCATTTGACAGTGAAGTTTGGATG 1616
Db GAGGTATATGATGGTACTGCTATTCAATCTGGCATTTGACAGTGAAGTTTGGATG 2310
QY 1617 TAC 1619
Db GAC 2313

RESULT 5

US-10-815-495-21

; Sequence 21, Application US/10815495

; Publication No. US20040191864A1

; GENERAL INFORMATION:

; APPLICANT: Novozymes Biotech, Inc.

; APPLICANT: Brody, Howard

; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient

; FILE REFERENCE: 10345-200-US

; CURRENT APPLICATION NUMBER: US/10/815,495

; CURRENT FILING DATE: 2004-03-31

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 21

; LENGTH: 2520

; TYPE: DNA

; ORGANISM: Aspergillus niger

; FEATURE:

```
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,c,g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: n=a,c,g or t
US-10-815-495-21

Query Match      52.4%; Score 908.2; DB 18; Length 2520;
Best Local Similarity 73.6%; Pred. No. 2.1e-270;
Matches 1537; Conservative 0; Mismatches 8; Indels 544; Gaps 8;

QY 1 TCACATCAAGCTCTCCCTTCTCTGAAATAAACCCCAAGAGGCAATTTATGATGGTCG 60
DB 151 TCACATCAAGCTCTCCCTTCTCTGAAATAAACCCCAAGAGGCAATTTATGATGGTCG 210
QY 61 CGTGTGTCTCTATTTCTGTACGGCTTCAGGTGCGGCACTGCTTTGGCTGCAACGC 120
DB 211 CGTGTGTCTCTATTTCTGTACGGCTTCAGGTGCGGCACTGCTTTGGCTGCAACGC 270
QY 121 CTGCGGACTGGGATCGCAATCATTTATTTCTCTACGGATCGATTTGCAAGACGG 180
DB 271 CTGCGGACTGGGATCGCAATCATTTATTTCTCTACGGATCGATTTGCAAGACGG 330
QY 181 ATGGGTGCGAGCTGCGACTTGTAACTGATGATGCGGATC----- 216
DB 331 ATGGGTGCGAGCTGCGACTTGTAACTGCGGATCAGGTGTGTTTACCTACTAGCTT 390
QY 217 -----AGAAATCTGTGTGGTGGAAACATGCGAGGCG 245
DB 391 TCAGAAAGAGGAATGTAAACTGACTGTGATAGAAATCTGTGTGGAAACATGCGAGGCG 450
QY 246 ATCATCGACA----- 255
DB 451 ATCATCGACAAGGTAAATTTGCCCTTTATCAAAAAAAGAGAAAGCAAGAA 510
QY 256 -----AGTTGGACTATATCCAGGGAATGGG 280
DB 511 TAAATAAAAAAGAACTAGTCTTAACCATCATAGTTGGACTATATCCAGGGAATGGG 570
QY 281 CTTTACAGGCATCTGATCACCCTGTTACAGCCGAGCTGCCGACACACCGCATATGG 340
DB 571 CTTTACAGGCATCTGATCACCCTGTTACAGCCGAGCTGCCGACACACCGCATATGG 630
QY 341 AGATGCTTACCATGCTTCTGCGAGCAGGATAT----- 373
DB 631 AGATGCTTACCATGCTTCTGCGAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC 690
QY 374 -----ATACTCTCTGAAACGAAAA 391
DB 691 TGTTCATCTTTTACATCAATATGAACATACTTGTATGGTTTTAGATACTCTCTGAAACGAAAA 750
QY 392 CTACGGCACTGAGATGATGTAAGCGCTCTCTTCCGCCCTTCTGAGAGGGGATGTA 451
DB 751 CTACGGCACTGAGATGATGTAAGCGCTCTCTTCCGCCCTTCTGAGAGGGGATGTA 810
QY 452 TCTTATGTCGATGTGTTGCTTAACCATAT----- 481
DB 811 TCTTATGTCGATGTGTTGCTTAACCATATGTTGTCGTGTCCTTTTGCAACTGACTTCGCG 870
QY 482 -----GGGCTATGATGGAGCGGTAGC 503
DB 871 GATATGGTTTCAATTCAGTACTGACATAGTAATATCAGGGCTATAGTGCAGGGGTAGC 930
QY 504 TCAGTCGATTAAGTGTGTTTAAACCGTTTCAGTTCCCAAGATCTATTCACCCGTTCTGT 563
DB 931 TCAGTCGATTAAGTGTGTTTAAACCGTTTCAGTTCCCAAGATCTATTCACCCGTTCTGT 990
QY 564 TTCATTCAAACTATCAAGATCAGACTCAGGTTGAGGATTTGCTGCTAGGATTAACACT 623
DB 991 TTCATTCAAACTATCAAGATCAGACTCAGGTTGAGGATTTGCTGCTAGGATTAACACT 1050
QY 624 GTCTCCTTGCCTGATCTCGATACCAACAGGATGTGTCGTCAGAAATGAAATGGTACGACTGG 683

1051 GTCTCCTTGCCTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGAATGGTACGACTGG 1110
684 GTGGGATCATTTGGTATCGAACTACTCCA----- 711
1111 GTGGGATCATTTGGTATCGAACTACTCCAAGTAGATATTTCTCCCTCATTTCTACAACCTGG 1170
712 -----TTGAGCGGCTCCGTATCGACACAGTAAAAACACTG 745
1171 CTGATCGATGATACTTACGAAATCAGTTTGACGGCTCCGTATCGACACAGTAAAAACACTG 1230
746 CCAGAAAGCACTTTCGCGCGGGTACAAACAAAGCCGAGCGGTGTAATCTGATCGGAGGT 805
1231 CCAGAAAGCACTTTCGCGCGGGTACAAACAAAGCCGAGCGGTGTAATCTGATCGGAGGT 1290
806 GCTCGACCGGTGATCCGGCTTACACTGTGTCCTTACAGAACGTCATGAGCGGCTACTGAA 865
1291 GCTCGACCGGTGATCCGGCTTACACTGTGTCCTTACAGAACGTCATGAGCGGCTACTGAA 1350
866 CTATCCCAT----- 874
1351 CTATCCCATGATGTTCTCTCAACCATGAGCTTCTTGCAGATCTCATCTCTTAACGAA 1410
875 -----TTATATCCACTCTCTCAAGCGCTTCAAGTCAACCTCCGCGAGCATG 920
1411 ACGGCTAAAAACCACTTACTTATCTCACTCTCAAGCGCTTCAAGTCAACCTCCGCGAGCATG 1470
921 GACGACCTCTCAACATGATCAACACCGTCAAACTCGGACTGTCCAGACTCAACACTCTCG 980
1471 GACGACCTCTCAACATGATCAACACCGTCAAACTCGGACTGTCCAGACTCAACACTCTCG 1530
981 GGCACATTTCTGCGAGAACCAACGACAAACCGGTTGCGTTT----- 1021
1531 GGCACATTTCTGCGAGAACCAACGACAAACCGTTCGTTTCTGTAAGTCTTCCCTTTTAT 1590
1022 -----TTACACCAACGACA 1035
1591 TTCCGTTTCCCAATTTCCACACAGAACCCCACTTAACAGAGCAAAAGTTACACCAACGACA 1650
1036 TAGCCCTCGCAAGAACGTCGCGAGCATTCATCATCTCAACGACGGAATCCCATCATCT 1095
1651 TAGCCCTCGCAAGAACGTCGCGAGCATTCATCATCTCAACGACGGAATCCCATCATCT 1710
1096 ACGCGCGCAAGAACGACATACGCGCGGCAAAACGACCCCGCAACCGCGAAGCAACCT 1155
1711 ACGCGCGCAAGAACGACATACGCGCGGCAAAACGACCCCGCAACCGCGAAGCAACCT 1770
1156 GGTCTCGGGCTTACCCGACCGGACGAGCTGTACAAGTTAATTCCTCCCGCAACGCA 1215
1771 GGTCTCGGGCTTACCCGACCGGACGAGCTGTACAAGTTAATTCCTCCCGCAACGCA 1830
1216 TCCGGAATGTCCTATGCAAGATACAGGATTCGTGACCTACA----- 1260
1831 TCCGGAATGTCCTATGCAAGATACAGGATTCGTGACCTACAAGSTAAAGCAACACT 1890
1261 ----- 1260
1891 CTAAGCATACCTAATGCGCTATCTTACAGATATCTGACACAAGAGACTAATCACTGGCA 1950
1261 -----AGAACTGCGCCATCTTACAAAGACGACAAACGATCGCCATCGCAAGGGCAAGAT 1316
1951 ATACAGAACTGCGCCATCTTACAAAGACGACAAACGATCCCGATCGCAAGGGCAAGAT 2010
1317 GGTCTCGAGATCGTCACTATCTTGTCCAAACAGGGTGTTCGGGTGATTCGTTATACCTTC 1376
2011 GGTCTCGAGATCGTCACTATCTTGTCCAAACAGGGTGTTCGGGTGATTCGTTATACCTTC 2070
1377 TCCTTGAGTGTGCGGGTTACACAGCGGCCAGCAATGACGAGGTCAATTCGGCTGCAG 1436
2071 TCCTTGAGTGTGCGGGTTACACAGCGGCCAGCAATGACGAGGTCAATTCGGCTGCAG 2130
1437 ACCGTGACGGTGTGTCGATGGAATGTGCTGTCCTATTCGCGAGGTGGGTACTACTAGG 1496
```


Db 2131 ACCGTGACGGTGGTTCCGATGGAATGTGCTGTCTTATGCGAGGTGGGCTACCTAGG 2190
QY 1497 GTATTGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTACT 1545
Db 2191 GTATTGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTACGGCT 2239

RESULT 6

US-10-653-047-3775
; Sequence 3775, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849,200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3775
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Aspergillus niger
US-10-653-047-3775

Query Match 46.5%; Score 806; DB 18; Length 1389;
Best Local Similarity 100.0%; Pred. No. 8,1e-239; Indels 0; Gaps 0;
Matches 806; Conservative 0; Mismatches 0;
QY 902 GTCAACTCCGGCAGATGAGACCTCTACAAATGATGATCAACCGTCAAAATCCGACTG 961
Db 1 GTCAACTCCGGCAGATGAGACCTCTACAAATGATGATCAACCGTCAAAATCCGACTG 60
QY 962 TCAGACTCAACACTCTCTGGGCACATTCGTGAGAACACGACCAACCCAGCGTTCCCTTC 1021
Db 61 TCAGACTCAACACTCTCTGGGCACATTCGTGAGAACACGACCAACCCAGCGTTCCCTTC 120
QY 1022 TTACACCAACGATAGCCCTCGCCAGAACGTCGAGCATTCATCTCTCAACGCGG 1081
Db 121 TTACACCAACGATAGCCCTCGCCAGAACGTCGAGCATTCATCTCTCAACGCGG 180
QY 1082 AATCCCATCATCTACCGCGGCCAAGAACAGCACTACGCGCGGAAAGACCCCGCGAA 1141
Db 181 AATCCCATCATCTACCGCGGCCAAGAACAGCACTACGCGCGGAAAGACCCCGCGAA 240
QY 1142 CCGCAAGCAACTGTGCTCTCGGGTACCCGACCGACGAGCTGTCAAGTTAATTGC 1201
Db 241 CCGCAAGCAACTGTGCTCTCGGGTACCCGACCGACGAGCTGTCAAGTTAATTGC 300
QY 1202 CTCGCGAAGCGAATCCGNACTATGCCATTAGCAAGATACAGGATTCTGTACCTACAA 1261
Db 301 CTCGCGAAGCGAATCCGNACTATGCCATTAGCAAGATACAGGATTCTGTACCTACAA 360
QY 1262 GAACTGGCCATCTCAAAAGACACAAACGATCGCATGCGCAAGGACACAGATGGGTC 1321
Db 361 GAACTGGCCATCTCAAAAGACACAAACGATCGCATGCGCAAGGACACAGATGGGTC 420
QY 1322 GCAGATCGTCACTATCTTGTCCAAAGGGTGTCTCGGGTGAATTCGTATACCTCTCTTT 1381
Db 421 GCAGATCGTCACTATCTTGTCCAAAGGGTGTCTCGGGTGAATTCGTATACCTCTCTTT 480
QY 1382 GAGTGTGGGGTTTACACAGCGCGCCAGCAATTGACGGAGGTCATTTGGCTGACAGCGGT 1441

Db 481 GAGTGTGGGGTTTACACAGCGCGCCAGCAATTGACGGAGGTCATTTGGCTGACAGCGGT 540
QY 1442 GACGGTTGGTTCGGATGGAATGTGCTGTCTTCTATGCGAGGTGGGCTACCTAGGTAAT 1501
Db 541 GACGGTTGGTTCGGATGGAATGTGCTGTCTTCTATGCGAGGTGGGCTACCTAGGTAAT 600
QY 1502 GTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGAG 1561
Db 601 GTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGAG 660
QY 1562 TATATGATGTTACTGCTATTTCAATCTGGCAATTTGGACAGTGAGTTTGTGATGTACAG 1621
Db 661 TATATGATGTTACTGCTATTTCAATCTGGCAATTTGGACAGTGAGTTTGTGATGTACAG 720
QY 1622 TTGGAGTGTGTTACTGCTGTCTATCTTATCTCTTCAATTTGTTTCGAACCTTAATGC 1681
Db 721 TTGGAGTGTGTTACTGCTGTCTATCTTATCTCTTCAATTTGTTTCGAACCTTAATGC 780
QY 1682 CAAGCAGCTAGTCTTATTTATAGGAA 1707
Db 781 CAAGCAGCTAGTCTTATTTATAGGAA 806

RESULT 7

US-10-213-990-41
; Sequence 41, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1488)
US-10-213-990-41

Query Match 45.3%; Score 786; DB 14; Length 1488;
Best Local Similarity 71.4%; Pred. No. 1.4e-232;
Matches 1035; Conservative 0; Mismatches 415; Indels 0; Gaps 0;
QY 94 TCGCGCACCTCTCTTGGCTGCAAGCCCTGCGAGCTGGCGATCGCAATCCATTTATTTCC 153
Db 38 TTGCTGGAACAGCTTTGGCGGCACCTCCAGCGCATGGCGATCCAGTCCATCTATTTCT 97
QY 154 TTCTCAGCGATCGATTTGCAAGGACGATGGGTGCGACGACTCGGACTTGTATATTCGG 213
Db 98 TACTCAGGATCGATTTGCAAGGACGATGGGTGCGACGACTCGGCTCGTGTGATAGTGT 157
QY 214 ATCAGAAATCTGTGTGGAACATGCGCGGGCATCTCCAGAGTTGGAGTATATTCAGG 273
Db 158 CTCGAGATGATTTGTGTGTGTATGTCAGAGGCAATTTTGAACAACTCGATTTACATCAAG 217
QY 274 GAATGGCTTTCACAGCATCTGATCACCCTCGTTACAGCCAGCTGCCCGACACCCG 333
Db 218 GATGGCTTTTACAGCAATTTGATTAACCCAGTCCAGCAAGCACTTCCCGAGTACGT 277
QY 334 CATATGGAGATGCTTACCATGGCTACTGGCAGCAGGATATATCTCTCTGAACAAACT 393
Db 278 CGAAGGCACTGCATACACCGGTACTGCGCAGCAAGCAATTTATTTCGATTCCTCAACT 337
QY 394 ACGGCACTGCAGATGATCTTGAAGGGCTCTCTCTCGGCCCTTCTATGAGAGGGGATGATTC 453

Db 338 ATGGACCGCTGACGACCTGAAAGCCCTGGCGTCAAGCTCTTCAAGACGGGCGCATGTATC 397
Qy 454 TTATGGTCGATGTGGTTGTAACCAATATGGGCTATGATGGAGCGGTAGCTCAGTCCGATT 513
Db 398 TCATGGTCGATGTGGTCATCATATGGGATATGCCGGTGCAGCGGATTCGGTCGACT 457
Qy 514 ACAGTGTGTTTAAACGGTTCAGTTCCTCAAGACTATCTCCACCCGTTCTGTTCATCAAA 573
Db 458 ATAGTGTCTTCAACCCCTTCAACTCCAGACCTCTTCCACCCATTTGCTTCATCAGCA 517
Qy 574 ACTATGAGATCAGACTCAGGTGTAGGATGCTGGCTAGGAGATAACACTGTCTCTTCG 633
Db 518 ATTACGATAATCAGACAGACGCTGGAATTTGCTGTGGGGACAATTCGGTTCCTTTAC 577
Qy 634 CTGATCTCGATACCAACCAAGGATGTGGTCAAGAAATGAATGGTACACTGGGTGGGATCAT 693
Db 578 CGATCTTGATACTACAAATCCGGATGTTCAAAAGATTTGGTACAAATTTGGGTGAACCTT 637
Qy 694 TGGTATCGAACTACTCCATTGACGGCTCCGTPATCGACACAGTAAACACGTCACGAAGG 753
Db 638 TAGTGTCTAAATTTCCATCGACGGCTACGAAATTGACACTGTGAAGCAGGTCCAGAGCG 697
Qy 754 ACTTCTGSCCGGGTACAACAAAGCGCAGGCGTGTACTGTATCGGCGAGGTGCTCGAGG 813
Db 698 ATTTCTGSCCGGGATTCAAATGACGCGCAGGCGTCTACTGTATCGGGAGGTATTGTGATG 757
Qy 814 GTGATCCGCGCTPACACTTGTCTCCCTACCAGAACGTCATGAGCGCGTACTGAACTATCCCA 873
Db 758 GAGATCCAGGTACACTTGTCTCCCTACCAGAGGTCCTGAGCGGGTGTGAAATACCCCTA 817
Qy 874 TTTACTATCCACTCTCAACGCGCTTCAAGTCAACCTCCGCGACGATGAGCAGCACTCTACA 933
Db 818 TATATTATCCGCTTCTGAAAGCCTTCCAAATCCACAAGTGGCAGTATGAGCAGCCTCTAAG 877
Qy 934 ACATGATCAACCGGTAAATCCGATGTTCAGACTCAACACTCTCTGGGCACTTCGTGCG 993
Db 878 ATATGATCAACCGGTCAAGTGCAGTGCAGTGCAGATTCGACGCTGTGTGGGCACTTTGTGCG 937
Qy 994 AGAAACAGACAAACCCAGGTTTCGCTCTTACCAACAGCATAGCCCTCGCCAGAGACG 1053
Db 938 AANAATCAGATACTCCCGGTTTCGATCGTATACCAAGACATAGCCCTCGCTAAAGACG 997
Qy 1054 TCGCAGCATTCATCATCTCTCAACGACGGAATCCCATCATCTACGCGCGCCCAAGAACAGC 1113
Db 998 CGCAGCATTCATCATCTCTCCGATGGAATCCCAATCATCTACGCGCGCCAGGAGCAAC 1057
Qy 1114 ACTAGCGCGGGGAAACGACCCCGGAAACCGGAAAGCAACTGTGGCTCTCGGGCTACGCGA 1173
Db 1058 ATTAAGCGCGGGCGGGATCCCGCAAAACCGGTGAGGCAAGTCTGGCTATCCGGCTACTCGA 1117
Qy 1174 CCGACAGCGAGGTGTACAAAGTTAATTTGCTCCGGAACGCAATCCGGAATATGCCATTA 1233
Db 1118 CAACAGGTGACCTATACAAAGTCTATCGGACAGCAAAACGCAATCCGGAATACGCGCATTA 1177
Qy 1234 GCAAGATACAGGATTCGTGACTTCAAGAACTGGCCCATCTACAAAGACGACACAAAG 1293
Db 1178 GCAAGAACCGGGATACGTGACTTATAGATTAACCCCATCTACAAAGATACCTCCACA 1237
Qy 1294 TCGCATCGCAAGGGCAAGATGGGTCCGAGATCGTGAATCTTTGTTCACAAAGGGTG 1353
Db 1238 TCGCATCGAAAGGGCTCCGACGAGCGCAGATCATCACCGTCTCTCGAACTCCGGTG 1297
Qy 1354 CTTCCGGGTGATTCGTATACCTCTCTTGAAGTGTGGGTTCAGACGCGCGCCAGGAT 1413
Db 1298 CTTCTGGAAGTCTTACACGCTCTCATTTGGGTGGAACAGGCTATGAGCGCGGACAAAC 1357
Qy 1414 TGACGAGGTCAATGGGTGACAGCGGTGAGCGGTGTTGCGATGCGAAATGTGCTGTTTC 1473
Db 1358 TGACTGAGATGTTCTCTTGACCAACCGGTGACCGTGGGCTCAGACAAAGGTTCGCCGTTT 1417
Qy 1474 CTATGGCAGGTGGGCTACCTAGGGTATTGTATCCGACTGAGAAGTTGGCAGGTAGCAAGA 1533

Db 1418 CCATGGCTAGTGGCTTGCCCGGGTGTTTTATCCCAACGGCTGGGCTGAACGGAAGTACTG 1477
Qy 1534 TCTGTAGTAG 1543
Db 1478 TTTGTACTTG 1487
RESULT 8
US-10-213-990-44
; Sequence 44, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1893)
US-10-213-990-44
Query Match 37.4%; Score 648.2; DB 14; Length 1893;
Best Local Similarity 65.4%; Pred. No. 9.4e-190;
Matches 966; Conservative 0; Mismatches 508; Indels 3; Gaps 1;
Qy 73 TATTTCGTAGCGCTTCAGGTCGCGCACCTGCTTTGGCTGCAACGCTCGCGACTGCG 132
Db 29 TGTCCCTGTGCTCGTCCCTGCTCGGACAGGCTGCCATGCTCTGACCCCGCAATGCG 88
Qy 133 GATCGCAATCCATTATTTCTTCTCAGGATCGATTTGCAAGACGGATGGTGCAGCA 192
Db 89 GCAGTCAATCGATCATTTCTCTGACCGGATCGGTTGCGCGGAGAACAAATTCACGA 148
Qy 193 CTGCGACTTGTAACTACGCGATCAGAAATACTGTGGTGGAAACATGGCAGGCACTATCG 252
Db 149 CTGCTGCTGGATGTACGCAACGACTGTATTCGGCGGAGCTGGCAGGGATCATCA 208
Qy 253 ACAAGTTGGACTATATCCAGGAAATGGGCTTTCACAGCCATCTGGATCACCCCGCTTACAG 312
Db 209 ATCATCTCGACTACATTCAAGGCAATGGATTTACTGCCATATGATCACCCCGCTAACTG 268
Qy 313 CCGAGTCCCGCAGACACCGCATATGGAGATGCTTACCATGGCTACTGGCAGCAGGATA 372
Db 269 AGCAGTCTTATGAGAACACCGCGCATGGTACTTTCGTACCATGGATACTGGCAGCAGATA 328
Qy 373 TATACTCTCTGAACGAAATCTACCGCACTGCAGATGACTTGAAGCGCTCTCTTCGGCCC 432
Db 329 TCCACGAGGTCAATGCCAATTTATGAAACGGCACAAGATCTTAGAGATCTGGCCAAACGCTC 388
Qy 433 TTCAATGAGAGGGGATGTATCTTATGGTCGATGTGGTTGCTTAACCATATATGGCTATGATG 492
Db 389 TGCACGCGGTGGGATGTACTTTGATGGTGCATGTGGTCGCAACCATATATGGGCTACAACG 448
Qy 493 GAGCGGTAGTCTAGTGCATTTACAGTGTGTTTAAACCGTTCAGTTTCCAAAGACTACTTCC 552
Db 449 GAGCGGAAATCTCGGTCAACTACCGTGTCTTCACTCCGTTTGAATTCGGCTACTTATTTCC 508
Qy 553 ACCGTTCTGTTTTCATTCAAAATCTGAAGATCAGACTCAGGTTTGAGGATTTGCTGGCTAG 612
Db 509 ACCACTACTGTCTCATCCGACTTACAAACCAACAGCTGTGGAGGACTGCTGGCTGG 568
Qy 613 GAGATAACACTGTCTCTCTGCTGATCTCGATATACCAACCAAGGATGTGTCACGAATGAAT 672

```
Db 659 GAGATACACTGCTCTGCTACCCGATCTCGACAGACGACGCGGAGTCCGAGCATCT 628
Qy 673 GGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTTGACGGCTCCGTTATCGACA 732
Db 629 GGTATGATTGGGTGAAGGATTTGGTTCGCAACTACTCCATCGACGCGCTCGCATCGACA 688
Qy 733 CAGTAAACACGTCGACAGAGACTTCTGGCCCGGGTACAAACGCGGAGGCGGTGACT 792
Db 689 CGGTGAAGCATGTGCGAAAGACTTCTGGCCCGGGTACAAATGACGCTGCTGGCGTCTACT 748
Qy 793 GTATCGCGAGGTGCTCGACGGGTGATCGGCGCTACACATTTGCTCCCTACCGAAGCTCATGG 852
Db 749 GTGTCGGTGAAGTCTTTTCGGGTGATCCACAATATACCTGTCCATACCAAGATTAACCTGG 808
Qy 853 ACGGCGTACTGAATCTCCCAATTTACTATCTCACTCTCAAGCGCTTCAAGTCAACCTCCG 912
Db 809 ATGGTGTACTCAACTACCCCATATACTATCAACTTCTCTACGGTTCCTCAATCGACCGAG 868
Qy 913 GCAGCATGACGACCTCTACACATGATCAACACCGTCAAAATCCGACTGTCCAGACTCAA 972
Db 869 GCAGCATGACCAATCTGTACCAATGATCAGCTCCGTTGGCTGTGACTGTGCGGATCCCA 928
Qy 973 CACTCTGGGSCATCTGTCGAGAACACGACACCCAGGTTGCTTCTTACACCAAG 1032
Db 929 CTTTGTCTGGCAACTTTATCGAAGACCATGATAACCCCGGATTTGCTCTATACGAGCG 988
Qy 1033 ACATAGCCCTCGCAAGAACGTCGACGATTTCAATCATCTCTCAACGACGGAATCCCATCA 1092
Db 989 ACTATTGCGAAGCAGAGAGCTCATCTCTTCAATGTTCTTCTCGAGCGCATCCCATTTG 1048
Qy 1093 TCTACCCCGGCAAGAACACGACTACGCGCGGGAACACGACCCGCGAACCAGCAAGCAA 1152
Db 1049 TCTACCCGCGACAGGAGACGACTACAGCGCGGTGCTGACCCCTGCGAACCCGAGGCTG 1108
Qy 1153 CTGGCTCTCGGCTACCGACGACGACGAGCTGTACAGTTAAATTTGCTTCGCGAAG 1212
Db 1109 TCTGGCTGTCTGGATACTCGACAGCGCTACGCTGTACAGCTGGATTGGCTCTACCAACA 1168
Qy 1213 CAATCGGAATATGCAATTAGCAAGATACAGGATTCGTGACTCAAGAACTGGGCCCA 1272
Db 1169 AGNTTCGCAAGCTAGGATTTCCAAAGACTCAGCTACATACATCCAGAACACCGCT 1228
Qy 1273 TCTCAAGACGACACAAACGATCGCATCGCGAAGGCGACA---GATGGTCCGACATCG 1329
Db 1229 TCTACTATGATTCCAATCTCTCGCTATGCGCAAGGCTCAGTCGCTGGCTCTCAAGTCA 1288
Qy 1330 TGACTATCTGTCCACAGGGTGTCTCGGGTATTCGTATACCTCTCTCTGAGTGGTG 1389
Db 1289 TTACCGTCTCAGTACAAAGGGATCCTCGGGCAGTTCTTACACCCCTCTCTCAGCGGCA 1348
Qy 1390 CGGGTTACACAGCGCGCCAGCAATTGACGGAGGTCAATGGCTGCACGACCGTGACGGTTG 1449
Db 1349 CGGGCTACTCTCGCGCGGCGCCACCTTGTGAGATGTATACATGCACTACTCTCAGCGTG 1408
Qy 1450 GTTCGGATGGAATGTGCTCTGTTCTTATGCGAGGTGGGCTACCTAGGGTATTTGATCCGA 1509
Db 1409 ACTCGAGCGGAATCTCGCGTGCCTCAATGATATCGGCTTGGCCAGAGTTTTCGTGCCCT 1468
Qy 1510 CTGAGAAGTGGCAGGTAGCAAGATCTGTAGTAGTCT 1546
Db 1469 CGTCATGGGTGCTGAGGAGTGGCTCTCGCGGCACTC 1505
```

RESULT 9

```
US-10-653-047-3775/c
; Sequence 3775, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
```

```
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3775
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Aspergillus niger
; US-10-653-047-3775
```

```
Query Match 25.3%; Score 439; DB 18; Length 1389;
Best Local Similarity 98.9%; Pred. No. 7e-125;
Matches 442; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1174 CCGACAGCGAGCTGTACAAAGTTAAATTTGCCCTCCGCGAACGCAATCCGGAATATGCCATTA 1233
Db 1364 CCGGCGTAGTGTGTACAAAGTTAAATTTGCCCTCCGCGAACGCAATCCGGAATATGCCATTA 1305
Qy 1234 GCAAAGATACAGGATTCGTGACCTCAAGAACTGGCCCATCTACAAGACGACACAACGA 1293
Db 1304 GCAAAGATACAGGATTCGTGACCTCAAGAACTGGCCCATCTACAAGACGACACAACGA 1245
Qy 1294 TCGCCATCGCAAGGCGACAGATGGTCCAGATCGTGTATCTTTGTCCAAAGGGTG 1353
Db 1244 TCGCCATCGCAAGGCGACAGATGGTCCAGATCGTGTATCTTTGTCCAAAGGGTG 1185
Qy 1354 CTTCCGGTGAATTCGTATACCTCTCTTCAGTGGTTCAGCGGTACAGCGCGGCGCAAT 1413
Db 1184 CTTCCGGTGAATTCGTATACCTCTCTTCAGTGGTTCAGCGGTACAGCGCGGCGCAAT 1125
Qy 1414 TGACGAGGTCAATTTGGCTGCACGACCGTTCAGTGGTTCGATGGAATGTCCCTGTTTC 1473
Db 1124 TGACGAGGTCAATTTGGCTGCACGACCGTTCAGTGGTTCGATGGAATGTCCCTGTTTC 1065
Qy 1474 CTATGCGAGTGGGCTACCTAGGGTATTTATCCGACTGAGAGTTGGCAGGTAGCAAGA 1533
Db 1064 CTATGCGAGTGGGCTACCTAGGGTATTTATCCGACTGAGAGTTGGCAGGTAGCAAGA 1005
Qy 1534 TCTGTAGTACCTCGTCAAGGTTGAGAGTATATGATGGTACTCTTCAATCTGGCATT 1593
Db 1004 TCTGTAGTACCTCGTCAAGGTTGAGAGTATATGATGGTACTCTTCAATCTGGCATT 945
Qy 1594 GGACAGTGAAGTTGAGTTTGAATGTACA 1620
Db 944 GGACAGTGAAGTTGAGTTTGAATGTACA 918
```

RESULT 10

```
US-10-653-047-3837
; Sequence 3837, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
```

```

; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3837
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(483)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3837

Query Match      21.6%; Score 374.2; DB 18; Length 483;
Best Local Similarity 91.0%; Pred. No. 5.1e-105;
Matches 443; Conservative 0; Mismatches 35; Indels 9; Gaps 4;

QY 1243 CAGGATTCGTGACCTACAAAGAACTGGCCCATCTACAAAGACGACACAAACGATCGCCATGC 1302
Db      |||
QY 1303 GCAAGGGCACAGATGGGTCGAGATCGTGACTATCTTGTCCAAACAAAGGCTCTCGGGTG 1362
Db      |||
QY 61 GCAAGGGCACAGATGGGTCGAGATCGTGACTATCTTGTCCAAACAAAGGCTCTCGGGTG 120
QY 1363 ATTCTGTATACCTCTCTTGTAGTGGTGGGTTACACAGCCGCCACGAAATTCAGCGAGG 1422
Db      |||
QY 121 ATTCTGTATACCTCTCTTGTAGTGGTGGGTTACACAGCCGCCACGAAATTCAGCGAGG 180
QY 1423 TCATTGGCTGCACGACCGGTGAGTGGTTCGGATGGAAATGTGCTGTTCCTATGGCGAG 1482
Db      |||
QY 181 TCATTGGCTGCACGACCGGTGAGTGGTTCGGATGGAAATGTGCTGTTCCTATGGCGAG 240
QY 1483 GTGGGCTACTAGGGTATTTGATTCGACTGAGAGTGGCAGGTGACGAGTCTGTAGTA 1542
Db      |||
QY 241 GTGGGCTACTAGGGTATTTGATTCGACTGAGAGTGGCAGGTGACGAGTCTGTAGTA 300
QY 1543 GCTCTGTAGGGTGGAGTATATGATGTGCTGCTATTCATCTGGCATTTGGACA--GT 1600
Db      |||
QY 301 GCTCTGTAGGGTGGAGTATATGATGTGCTGCTATTCATCTGGCATTTGGACA--GT 359
QY 1601 GAGTTTGAGTTTGATGTACAGTTGGAGTGGTGTACTGTCTGCTATCCCTTTATCTTCGA 1660
Db      |||
QY 360 AAGTTTGAGTTTGATGTACAGTTGGAGTGGTGTACTGTCTGCTATCCCTTTATCTTCGA 418
QY 1661 TTGTTTTCGAAACCTTAATGCCAAACGCTAGTCTATTTATAGGAAAAAAGAAAAA 1720
Db      |||
QY 419 ----GTTTTCGAAACCTTAATGCCAAACGCTAGTCTATTTATAGGAAAAAGGNTCCGGATTA 473
QY 1721 AAAAAA 1727
Db      |||
QY 474 AAAAAA 480
Db      |||

RESULT 11
US-10-128-590-28/c
; Sequence 28, Application US/10128590
; Publication No. US20030108561A1
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Core-glycosylated HCV envelope proteins
; FILE REFERENCE: 135 PCT
; CURRENT APPLICATION NUMBER: US/10/128,590
; CURRENT FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4190
; TYPE: DNA
; ORGANISM: vector pMF30
US-10-128-590-28
```

```

Query Match      21.3%; Score 369.2; DB 15; Length 4190;
Best Local Similarity 54.6%; Pred. No. 6e-103;
Matches 759; Conservative 0; Mismatches 628; Indels 3; Gaps 1;

QY 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCCCTCTCACGGATCGATTTGCAAGACGG 180
Db      |||
QY 1671 CTGATAAATGGAAGACCAATCGATTATCAAAATCGTTACTGATAGGTTGCGAGATCG 1612
QY 181 ATGGGTCGACGACTCGGACTTGTATATACTGCGGATCAGAAATACCTGTGTGGAACATG 240
Db      |||
QY 1611 ATGGGTCGACGACTCGGACTTGTATATACTGCGGATCAGAAATACCTGTGTGGAACAT 1552
QY 241 AGGGCATCATCGACAAGTTGGACTATATCCAGGGAAATGGGCTTTCACAGCCATCTGGAT 300
Db      |||
QY 1551 AAGGATTTATCGACAAGTTGGATTTATTTCAAGTATGCGTTTCACTGCGATCTGGATCT 1492
QY 301 CCCCCGTTACAGCCAGCTGCCCCAGACACCGCATATGAGAGATGCTTACCATGGTACT 360
Db      |||
QY 1491 CCCCAGTTGTTGAGCAAAATTCCTGACAAATCTGCTTATGCTTATGCTTACCATGGTTAT 1432
QY 361 GGCACAGGATATATATCTCTGACGAAACCTACGGCACTCGAGATGACTTGAAGCGC 420
Db      |||
QY 1431 GGATCAAAATATATGATGAATTTGAACACTAATTTTGGTACCGCTGATGAATTTGAACAAT 1372
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTTATGGTCGATGTGGTTGCTTAACCAT 480
Db      |||
QY 1371 TAGCTAGCGAATTCATCCAGAACATGTTATGATGGTCGACGTTGTTTACACCAT 1312
QY 481 TGGGCTATGATGAGCGGGTAGCTCAGTCGATTAAGTGTGTTTAAACCGTTCAAGTTCCC 540
Db      |||
QY 1311 ATGCTTGGAAACGAGATGGTTCAAGCGTAGATTTATTTAGTTTTCATCTCCATCAATCA 1252
QY 541 AAGACTACTTCCACCGTTCTGTTTCATTTCAAACTATGAAGTACAGACTCAGGTTGAG 600
Db      |||
QY 1251 AATCTTACTTCCAGATTTATTTGTTGATTAACAAATTAATGATCAAACTAATGTTGAAG 1192
QY 601 ATTGCTGGCTAGGAGATAACACTGTCTCTCTCCCTGATCTCGATATACCAAGGATGTGG 660
Db      |||
QY 1191 ATTGTTGGAAGGTGATCTGAACTCTCCCTCCAGATTTAAGTACCGAGGATATGAAG 1132
QY 661 TCAAGAAATGAATGGTACGACTGGGTGGGATCAATGGTATCGAACTACTCCATTTGAGCGGC 720
Db      |||
QY 1131 TTATAGGAGTATTTCAAACCTTGGGTGTGAGATTTTGTTCAAAACCTATTTCAATCGATGG 1072
QY 721 TCCGCTATCGACACAGTAAACACAGTCCAGAGAGACTTCTGGCCCGGGTACAAAGACCG 780
Db      |||
QY 1071 TAAGAAATGATAGTGCAAGACGACGTAGATACCGCTTTCATTAACGAAGTTTGAGGACGCT 1012
QY 781 CAGGCGTGTACTGTATCGGCGAGGTGCTCGACGGTGTATCCGGCCTACACTTGTCCCTACC 840
Db      |||
QY 1011 CTGGTGTATTAACCTTAGGTGAAGTTTATCAAGGAGATCCAACTTATCTTGTCCATATC 952
QY 841 AGAAGCTCATGGAACGGGCTGTAAGACTATCCCATTTATCTATCCACTCTCCCAAGCCCTCA 900
Db      |||
QY 951 AGAATTTATGAAAGAGGTTTACCAACTATCCATTTACTATCTATCCAGTATATAGATTTCT 892
QY 901 AGTCAACCTCCGGCAGCATGGACAGCTCTACAAATGATCAACACCGTCAAACTCCGACT 960
Db      |||
QY 891 GTGATACCTTCGGCGACTTCCAGTGAGTTAACTTCAATGATCTCCAGGTTACAGTCACTCT 832
QY 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAAACCAACCAACCCACCGTTCCGCTT 1020
Db      |||
QY 831 GTTCGAGCGTCTCTTTGTTGGAAACTTTATTTGAAACCAATGACCCAGTTAGATTTCCAT 772
QY 1021 CTTACACCAACGACATAGCCCTCGCCAAAGAGCTGCGAGCATTCATCTCTCAAGACG 1080
Db      |||
QY 771 CAGTTACCTCGACACATCTTTGATTAAGATGACATGGCTTTTATAATTTTGGGTGATG 712
QY 1081 GAATCCCATCATCTACGCCGCCAGACAGCATAGCCGCGGAAACGACCCCGCA 1140
Db      |||
QY 711 GTATCCCAATTTATTTATGCGCCAAAGCAAGGCTCTCAATGGTGTCCGATCTCGCCA 652
```



```
QY 1501 TGTATCCGAC 1510
Db 294 ACGTCCGTC 285

RESULT 13
US-10-128-578B-28/c
; Sequence 28, Application US/10128578B
; Publication No. US20030211597A1
; GENERAL INFORMATION:
; APPLICANT: BOSMAN, FONS
; APPLICANT: DEPLA, ERIK
; APPLICANT: DESCHAMPS, GEERT
; APPLICANT: SABLON, ERWIN
; APPLICANT: SAMSON, ISABELLE
; APPLICANT: VAN BROEKHOVEN, ANNIE
; APPLICANT: HAELEWYN, JOOST
; TITLE OF INVENTION: EXPRESSION OF CORE-GLYCOSYLATED HCV ENVELOPE
; FILE REFERENCE: 2551-73
; CURRENT APPLICATION NUMBER: US/10/128,578B
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 2.1
; SEQ ID NO 28
; LENGTH: 4190
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Vector pMP30
US-10-128-578B-28

Query Match 21.3%; Score 369.2; DB 17; Length 4190;
Best Local Similarity 54.6%; Pred. No. 66-103; Indels 3; Gaps 1;
Matches 759; Conservative 0; Mismatches 628;

QY 121 CTGGGACTGGGGATCGCAATCATTTATTTCTCTACGGATCGATTGTCAGAGCGG 180
Db 1671 CTGATAAATGGAAGACCAATCGATTATCAATCGTTACTAGTAGTTGGCAGATCTG 1612

QY 181 ATGGGTCGACGACTGGCATCTGTATATCTGCGGATCAGAAATCTGTGTGGAACATGGC 240
Db 1611 ATGGGTCGACGACGACTGTGTTAGTGAGTGATGTCGCAAGTACTGTGGTGATCTTATA 1552

QY 241 AGGGATCATCAGAGTTGGACTATATCCAGGATGGCTTCACGCCATCTCGATCA 300
Db 1551 AAGGGATTATCGACAAGTTGGATTATATCAAGGTATGGGTTTCACTGGCATCTGATCT 1492

QY 301 CCCCCGTTACAGCCGACTGCCCGACACCGCATATGGAGATGCTTACCATGGCTACT 360
Db 1491 CCCCAGTTGTTGAGCAAAATCTCGACAATCTGCTTATGTTATGCTTACCATGGTTATT 1432

QY 361 GGACAGAGATATATCTCTCTGAACGAAAATCTACGGCACTGCGAGATGACTGAAGCGC 420
Db 1431 GGATGAAAATATTTGATGAATTGAACACTAAATTTTGGTACCGCTGATGAATTTGAAAAT 1372

QY 421 TCTCTCGGCCCTTCATGAGAGGGGATGATCTTATGCTGCTGATGCTGCTGCTACCAT 480
Db 1371 TAGCTAGCGAAATGCAATTCAGAAAGCATGTTATGATGCTGACGCTGTTTCAACCAATT 1312

QY 481 TGGGCTATGATGGAGCGGCTAGCTCAGTCGATATACAGTGTGTTAAACCGTTTCAGTCC 540
Db 1311 ATGCTTGGAAACGAGATGGTTCAAGGTAGATTATTTCTAGTTTCACTCCATCAATCAAC 1252

QY 541 AAGACTACTTCCACCGCTTCTGTTTCATTTCAAACTATCAAGATCAGATCAGGTTGAGG 600
Db 1251 AATCTTACTTCCACGATTATTTGTTGATTACAAATTAATGATCAACCAATGTTGAAG 1192

QY 601 ATTGCTGGCTAGGATATACATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 1191 ATTGTTGGGAAGGTGATCTGAAGTCTCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
```

```
661 TCAGAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAATCTACTCTCATTTGACGGC 720
1131 TTATAGGAGTATTTCAAACCTTGGGTGTGAGATTTTGTTCAAAACCTATTCAATCGATGTT 1072
721 TCCGTATCGACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGGTACAAACAAGCG 780
1071 TAAGAAATTGATAGTCNAAGCACGATAGATACCGCTTCATTAAACGAAGTTTGGAGCGCTT 1012
781 CAGCGCTGTACTGTATCGCGAGGTGCTCGAGCGGTGATCCGGCCTACACTTGTCCCTACC 840
1011 CTGGTGTATTAACCTTAGTGAAGTTTATCAAGGAGATCCAACTTATATCTTGTCCATATC 952
841 AGAAGCTCATGACCGCGTACTGAACTATCCCATTTATCTATCCACTCTCTCAAGCGCTTCA 900
951 AGAATTATATGAAGGAGTTACCAACTATCCATTTATCTATCCAGTATATAGATTTCTTCA 892
901 AGTCAACCTCCGGCAGCATGGAGCCTCTCAACATGATCAACACCGTCAATCCGACT 960
891 GTGATACTTCGGCGACTTCCAGTGAGTTAACTTCAATGATCTCCACGTTACAGTCACTTT 832
961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGCGAACCACGACCAACACCGTTCGCTT 1020
831 GTTCGAGCTCTCTTGTGTTGGAAACTTTATTTGAAGAACCATGACCAAGTTAGATTTCCAT 772
1021 CTTACACCAACGACATAGCCCTCGCCAAAGACGTGCGAGCATTCATCTCTCAAGCAG 1080
771 CAGTTACCTCGACACATCTTGAATTAAGATGACATGGCTTTTATAATTTTGGGTGATG 712
1081 GAATCCCCCATCTACCGCGCGCAAGACACGACTACGCGCGCGGGAACGACCGCGCA 1140
711 GTATCCCAATTTATTTATGCGCCAAAGAACAAAGGCTCTCAATGGTGGTTCGGATCTCGCA 652
1141 ACCGGAAGCAACCTCGGCTCTCGGCTACCCGACGACGAGCGCTGTACAAAGTTAAATTG 1200
651 ATAGAGAGCTTTATGTTTGTAGTGATATATACCGATTCAGAACTACTACGAGCTAATCA 592
1201 CTTCCGGAACCAATCCGGAATATGCAATTTAGCAATAGCAAGATTCGTTGACCTACA 1260
591 GTPAACTAAATCAATAAGAAATCAAGCTATTTAAGAAAGGATTCGCTATTCAACTTACA 532
1261 AGAATCGGCCCATCTACAAAGACGACACACGATCGCCATGCGCAAGGCGCAGAGCGT 1320
531 AATCCTCAGTTGTTTCTTCTTCAGACCACTATATAGCCACTATAGGAAGGATGCGGATG 475
1321 CCGCATGCTGACTATCTTGTCCAAAGGGGTCTTCGGGTGATTCGTTATACCTCTCCT 1380
474 CTAATCAATGATTTCATTTTATATATTTAGTTTCAACCGGTCACAGGATATTACTG 415
1381 TGAGTGGTGGGGTTTACACAGCCGCCAGCAATTCACGGAGGTCAITGGCTGACGACCG 1440
414 TCAGCAACACCGGCTATTCTAGTGGTGATAAAGTTATCGATATTATTTCTTGCATTCGG 355
1441 TCACGGTGGTTCGAGTGGAAATGTCCTGTTCTTATGCGAGGTGGCTTACCTAGGATAT 1500
354 TTTTACGCTGGTACCTCCGGAAGCTTATCTGATCAATTTCTGGTGAATGCGCAAGATTT 295
1501 TGTATCCGAC 1510
294 ACGTCCGTC 285
```

RESULT 14
US-10-653-047-3796
; Sequence 3796, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene

```
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849,200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3796
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(605)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3796

Query Match      21.3%; Score 368.8; DB 18; Length 605;
Best Local Similarity 97.2%; Pred. No. 2.7e-103;
Matches 418; Conservative 0; Mismatches 7; Indels 5; Gaps 4;

QY 96 GCGGCACCTGCTTGGCTGCAAGCGCTGCGGACTGCGGATCGCAATCCATTTATTTCTT 155
Db 1 GCGGCACCTGCTTGGCTGCAAGCGCTGCGGACTGCGGATCGCAATCCATTTATTTCTT 60

QY 156 CTCACGGATCGATTTTCAAGGACGGATGGTTCGACGACTGCGACTTGTAACTACTGCGAT 215
Db 61 CTCACGGATCGATTTTCAAGGACGGATGGTTCGACGACTGCGACTTGTAACTACTGCGAT 120

QY 216 CAGAAATACTGTGGTGGAAATGCGCAGGACATCATCGAAGTTGCAATATATCCAGGGA 275
Db 121 CAGAAATACTGTGGTGGAAATGCGCAGGACATCATCGAAGTTGCAATATATCCAGGGA 180

QY 276 ATGGGCTTCACAGCCATCTGGATCCACCGCTTACAGCCAGCTGCCCGCAGACCCGCA 335
Db 181 ATGGGCTTCACAGCCATCTGGATCCACCGCTTACAGCCAGCTGCCCGCAGACCCGCA 240

QY 336 TATGAGATGCTTACCATGGCTACTGGCAGCAGGATATATCTCTGAAACGAAACTAC 395
Db 241 TATGAGATGCTTACCATGGCTACTGGCAGCAGGATATATCTCTGAAACGAAACTAC 300

QY 396 GGCATGCGAGATGACTTGAAGCGCTCTTTTGGGCCCTTTCATGAGAGGGGATGTATCTT 455
Db 301 GGCATGCGAGATGACTTGAAGCGCTCTTTTGGGCCCTTTCATGAGAGGGGATGTATCTT 358

QY 456 ATGGTTCGATGCTGGTTCCTAACCATATGGGCTATGATGGAGCGGGTAGCTCAGTCGAT 515
Db 359 ATGGTTCGATGCTGGTTCCTAACCATATGGGCTATGATGGAGCGGGTAGCT-AGTCGAT 415

QY 516 AGTGTGTTTA 525
Db 416 AGTGTGTTAA 425

RESULT 15
US-10-653-047-3776
; Sequence 3776, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849,200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
```

```
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3776
; LENGTH: 1048
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1048)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3776

Query Match      19.9%; Score 344.2; DB 18; Length 1048;
Best Local Similarity 67.7%; Pred. No. 1.6e-95;
Matches 481; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 TACTGTATCGCGAGGTGCTCGACGGTGATTCGGGCTACACTTGTCCCTACCAAGACGTC 848
Db 3 TGCTGCTCGCGGAAAGTCGACAACCGCAACCTCGCCTCGACTGCCCACATACCAAGAGTGC 62

QY 849 ATGGACGGGCTACTGAACTATCCCATTTACTATCCACTCTCAACGGCTTCAAGTCAACC 908
Db 63 CTGGACGGGCTCTCAACTATCCGATCTACTGGCAACTCTCTACGGCTTCGAATCCTCC 122

QY 909 TCCGGCAGCATGGACGACCTCTACCAACATGATCAACACCGTCAAAATCCGACTTCCAGAC 968
Db 123 ACGGCAGCATCAGCAACCTCTCAACATGATCAAAATCCGTCGCAAGCACTGCTCGAT 182

QY 969 TCAACACTCTCTGGGCACATTTGTCGAGAACACGACCAACCCACGGTTCGCTTCTTACCC 1028
Db 183 CGACACTACTCGGCAACTTTCATCGAAAACACGACAAATCCCGTTTCGCTCTACACC 242

QY 1029 AACGACATAGCCCTCGCCCAAGACGTCGACGATTCATCATCTCAACACGGAATCCCC 1088
Db 243 TCCGACTACTTCGACAGCCAAACGCTCTCAGCTACATCTTCTCTCCAGCGCATCCCC 302

QY 1089 ATCATCTACGCGCGCAAGAACAGACACTACGCGCGGGAAGACGACCCCGGAAACCGGAA 1148
Db 303 ATCTCTACGCGCGGGAAGAACAGACACTACTCTCGGGGCAAGGTGCGCTTACACCCGGA 382

QY 1149 GCAACCTGCTCTCGGGCTACCCGACCGACAGCTGTACAAGTTAATTCCTCCGCG 1208
Db 363 GCGACTTGGCTTTCAGGCTACGACACCTCCGACAGCTGTACACCTGGATAGCCACACG 422

QY 1209 AACGCAATCCGGAATGTCCTATGCCATTTAGCAAGATACAGGATTCGTGACCTACAAGAACTGG 1268
Db 423 AACGGGATCCGCAAACTAGCCATCTCAGCTGACTCGGNCCTACATTAACCTACGCGAATGAT 482

QY 1269 CCCATCTACAAAGACGACACACAGATCGCATCGCAAGGGACACAGATGGTTCGACAGATC 1328
Db 483 GCATTTACATGACAGCAACACCATCGCAATGCGCAAGGACACCTCAGGAGGCCAAGTC 542

QY 1329 GTGACTATCTTGTCCAACAAGGGTCTTCGGGTGATTCGTATACCTCTCTTGTAGTGT 1388
Db 543 ATCAGCGTCTCTCAACAAGGCTCTTCAGGAAGCAGCTACACCTCGACCTCAGCGGA 602

QY 1389 GCGGGTTACAGCGCGGCGCAGCAATTTAGCGAGGTGATTCGGCTGACAGCCGTCGAGCGTT 1448
Db 603 AGCGGCTACATATCCGCGCAAGAGTGTGATCGAAGCGTACACATGCAATCCGTCGACCGTG 662

QY 1449 GGTTCGGATGGAAGTGCCTGTTCTTCTATGGCAGGTGGGCTACCTAGGT 1498
Db 663 GACTCAGCGCGGATATTCCCGTGGATGGCGTTCGGGATTCGGAGAGT 712

Search completed: February 12, 2005, 04:03:50
Job time : 960 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2005, 19:15:18 ; Search time 6083 Seconds
(without alignments)
11636.883 Million cell updates/sec

Title: US-10-820-200-1

Perfect score: 1734

Sequence: 1 teacatcaagctctcccttc.....aaaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Parents NA Main:*

- 1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq:*
- 3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq:*
- 4: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
- 5: /cgn2_6/ptodata/1/pna/US07 COMB.seq:*
- 6: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
- 7: /cgn2_6/ptodata/1/pna/US081 COMB.seq:*
- 8: /cgn2_6/ptodata/1/pna/US082 COMB.seq:*
- 9: /cgn2_6/ptodata/1/pna/US083 COMB.seq:*
- 10: /cgn2_6/ptodata/1/pna/US084 COMB.seq:*
- 11: /cgn2_6/ptodata/1/pna/US085 COMB.seq:*
- 12: /cgn2_6/ptodata/1/pna/US086 COMB.seq:*
- 13: /cgn2_6/ptodata/1/pna/US087 COMB.seq:*
- 14: /cgn2_6/ptodata/1/pna/US088 COMB.seq:*
- 15: /cgn2_6/ptodata/1/pna/US089 COMB.seq:*
- 16: /cgn2_6/ptodata/1/pna/US090 COMB.seq:*
- 17: /cgn2_6/ptodata/1/pna/US091 COMB.seq:*
- 18: /cgn2_6/ptodata/1/pna/US092A COMB.seq:*
- 19: /cgn2_6/ptodata/1/pna/US092B COMB.seq:*
- 20: /cgn2_6/ptodata/1/pna/US093A COMB.seq:*
- 21: /cgn2_6/ptodata/1/pna/US093B COMB.seq:*
- 22: /cgn2_6/ptodata/1/pna/US094 COMB.seq:*
- 23: /cgn2_6/ptodata/1/pna/US095A COMB.seq:*
- 24: /cgn2_6/ptodata/1/pna/US095B COMB.seq:*
- 25: /cgn2_6/ptodata/1/pna/US095C COMB.seq:*
- 26: /cgn2_6/ptodata/1/pna/US095D COMB.seq:*
- 27: /cgn2_6/ptodata/1/pna/US096A COMB.seq:*
- 28: /cgn2_6/ptodata/1/pna/US096B COMB.seq:*
- 29: /cgn2_6/ptodata/1/pna/US096C COMB.seq:*
- 30: /cgn2_6/ptodata/1/pna/US096D COMB.seq:*
- 31: /cgn2_6/ptodata/1/pna/US096E COMB.seq:*
- 32: /cgn2_6/ptodata/1/pna/US097A COMB.seq:*
- 33: /cgn2_6/ptodata/1/pna/US097B COMB.seq:*
- 34: /cgn2_6/ptodata/1/pna/US097C COMB.seq:*
- 35: /cgn2_6/ptodata/1/pna/US098A COMB.seq:*
- 36: /cgn2_6/ptodata/1/pna/US098B COMB.seq:*
- 37: /cgn2_6/ptodata/1/pna/US098C COMB.seq:*
- 38: /cgn2_6/ptodata/1/pna/US098D COMB.seq:*
- 39: /cgn2_6/ptodata/1/pna/US099A COMB.seq:*
- 40: /cgn2_6/ptodata/1/pna/US099B COMB.seq:*
- 41: /cgn2_6/ptodata/1/pna/US099C COMB.seq:*
- 42: /cgn2_6/ptodata/1/pna/US099D COMB.seq:*
- 43: /cgn2_6/ptodata/1/pna/US099E COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US099G COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US100A COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US100B COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US101A COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US101B COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US102A COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US102B COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US103A COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US103B COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US104A COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US104B COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US105A COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US105B COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US106A COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US107A COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US107B COMB.seq:*

61: /cgn2_6/ptodata/1/pna/US107C COMB.seq:*

62: /cgn2_6/ptodata/1/pna/US107D COMB.seq:*

63: /cgn2_6/ptodata/1/pna/US108A COMB.seq:*

64: /cgn2_6/ptodata/1/pna/US108B COMB.seq:*

65: /cgn2_6/ptodata/1/pna/US109A COMB.seq:*

66: /cgn2_6/ptodata/1/pna/US109B COMB.seq:*

67: /cgn2_6/ptodata/1/pna/US109C COMB.seq:*

68: /cgn2_6/ptodata/1/pna/US110 COMB.seq:*

69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq:*

70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq:*

71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq:*

72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq:*

73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq:*

74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq:*

75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq:*

76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq:*

77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq:*

78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq:*

79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq:*

80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq:*

81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq:*

82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq:*

83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq:*

84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq:*

85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq:*

86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq:*

87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq:*

88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq:*

89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq:*

90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq:*

91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq:*

92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq:*

93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq:*

94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq:*

95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq:*

96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq:*

97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq:*

98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq:*

99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq:*

100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq:*

101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq:*

103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq:*

104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq:*

105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq:*

106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq:*

107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq:*

108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq:*

109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq:*

110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq:*

111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq:*

112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq:*

113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq:*

114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq:*

115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq:*

116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6047 COMB.seq.*
118: /cgn2_6/ptodata/1/pna/US6048 COMB.seq.*
119: /cgn2_6/ptodata/1/pna/US6049 COMB.seq.*
120: /cgn2_6/ptodata/1/pna/US6050 COMB.seq.*
121: /cgn2_6/ptodata/1/pna/US6051 COMB.seq.*
122: /cgn2_6/ptodata/1/pna/US6052 COMB.seq.*
123: /cgn2_6/ptodata/1/pna/US6053 COMB.seq.*
124: /cgn2_6/ptodata/1/pna/US6054 COMB.seq.*
125: /cgn2_6/ptodata/1/pna/US6055 COMB.seq.*
126: /cgn2_6/ptodata/1/pna/US6056 COMB.seq.*
127: /cgn2_6/ptodata/1/pna/US6057 COMB.seq.*
128: /cgn2_6/ptodata/1/pna/US6058 COMB.seq.*
129: /cgn2_6/ptodata/1/pna/US6059 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
------------	-------	-------------	--------	----	----	-------------

1	1734	100.0	1734	32	US-09-710-339-1	Sequence 1, Appli
2	1734	100.0	1734	32	US-09-710-339A-1	Sequence 1, Appli
3	1734	100.0	1734	32	US-09-710-339B-1	Sequence 1, Appli
4	1734	100.0	1734	63	US-10-820-200-1	Sequence 1, Appli
5	1707	98.4	1914	24	US-09-533-559-5355	Sequence 5355, Ap
6	1707	98.4	1914	58	US-10-653-047-5355	Sequence 5355, Ap
7	1497.8	86.4	1827	2	PCT-US04-20499-39	Sequence 39, Appl
8	1497.8	86.4	1827	64	US-10-877-849-39	Sequence 39, Appl
9	1497.4	86.4	1923	2	PCT-US04-20499-35	Sequence 35, Appl
10	1497.4	86.4	1923	64	US-10-877-849-35	Sequence 35, Appl
11	1011.2	58.3	3285	2	PCT-US02-27129-46	Sequence 46, Appl
12	1011.2	58.3	3285	2	PCT-US04-07182A-46	Sequence 46, Appl
13	1011.2	58.3	3285	50	US-10-228-063-46	Sequence 46, Appl
14	979	56.5	2443	2	PCT-US04-10126-17	Sequence 17, Appl
15	979	56.5	2443	63	US-10-815-495-17	Sequence 17, Appl
16	979	56.5	2443	115	US-60-459-902-17	Sequence 17, Appl
17	908.2	52.4	2520	2	PCT-US04-10126-21	Sequence 21, Appl
18	908.2	52.4	2520	63	US-10-815-495-21	Sequence 21, Appl
19	908.2	52.4	2520	115	US-60-459-902-21	Sequence 21, Appl
20	806	46.5	1389	24	US-09-533-559-3775	Sequence 3775, Ap
21	806	46.5	1389	58	US-10-653-047-3775	Sequence 3775, Ap
22	786	45.3	1488	1	PCT-US02-24842-41	Sequence 41, Appl
23	786	45.3	1488	50	US-10-213-990-41	Sequence 41, Appl
24	786	45.3	1488	100	US-60-309-870-41	Sequence 41, Appl
25	759.8	43.8	1485	55	US-10-486-868-7	Sequence 7, Appli
26	675.2	38.9	1533	2	PCT-US04-20499-7	Sequence 7, Appli
27	675.2	38.9	1533	64	US-10-877-849-7	Sequence 7, Appli
28	675.2	38.9	1830	2	PCT-US04-20499-37	Sequence 37, Appl
29	675.2	38.9	1830	64	US-10-877-849-37	Sequence 37, Appl
30	675.2	38.9	1890	2	PCT-US04-20499-33	Sequence 33, Appl
31	675.2	38.9	1890	64	US-10-877-849-33	Sequence 33, Appl
32	675.2	38.9	1923	2	PCT-US04-20499-31	Sequence 31, Appl
33	675.2	38.9	1923	64	US-10-877-849-31	Sequence 31, Appl
34	648.2	37.4	1893	1	PCT-US02-24842-44	Sequence 44, Appl
35	648.2	37.4	1893	50	US-10-213-990-44	Sequence 44, Appl
36	648.2	37.4	1893	100	US-60-309-870-44	Sequence 44, Appl
37	439	25.3	1389	24	US-09-533-559-3775	Sequence 3775, Ap
38	439	25.3	1389	58	US-10-653-047-3775	Sequence 3775, Ap
39	382.4	22.1	1479	2	PCT-US04-07096-498	Sequence 498, App
40	375.4	21.6	1404	10	US-08-467-831A-7	Sequence 7, Appli
41	375.4	21.6	1404	10	US-08-470-702B-7	Sequence 7, Appli
42	374.2	21.6	483	24	US-09-533-559-3837	Sequence 3837, Ap
43	374.2	21.6	483	58	US-10-653-047-3837	Sequence 3837, Ap
44	373.8	21.6	1404	10	US-08-467-831A-1	Sequence 1, Appli
45	373.8	21.6	1404	10	US-08-467-831A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-710-339-1
; Sequence 1, Application US/09710339
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Pedersen, Sven
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Fungamyl-like Alpha-Amylase Variants
; FILE REFERENCE: 5835.200-US
; CURRENT APPLICATION NUMBER: US/09/710,339
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/165,786
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Aspergillus Oryzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(1547)
; OTHER INFORMATION: mat_peptide
US-09-710-339-1
Query Match 100.0%; Score 1734; DB 32; Length 1734;
Best Local Similarity 100.0%; Pred. No. 7.8e-252;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACATCAAGCTCTCCCTTCTCTGAAACAATAAACCCACAGAGGAGGATTTATGATGGTGC 60
Db 1 TCACATCAAGCTCTCCCTTCTCTGAAACAATAAACCCACAGAGGAGGATTTATGATGGTGC 60
QY 61 CTGTGTGTCTTATTTCTGTACGGCTTCAGGTCGGCGACCTCTTGGCTGCAACGC 120
Db 61 CTGTGTGTCTTATTTCTGTACGGCTTCAGGTCGGCGACCTCTTGGCTGCAACGC 120
QY 121 CTGCGGACTGGCGATCGCAATCATTTATTTCTTCTCAGGATCGATTTGCAAGGACGC 180
Db 121 CTGCGGACTGGCGATCGCAATCATTTATTTCTTCTCAGGATCGATTTGCAAGGACGC 180
QY 181 ATGGTTCGACGACTCGGACTTGTAAATATATGCGGATCGAATAATATGTTGGTGAACATGCG 240
Db 181 ATGGTTCGACGACTCGGACTTGTAAATATATGCGGATCGAATAATATGTTGGTGAACATGCG 240
QY 241 AGGGCATCATGACAAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCATCTGGATCA 300
Db 241 AGGGCATCATGACAAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCATCTGGATCA 300
QY 301 CCCCCTTTACAGCCCGAGCTGCCCGACAGACCCAGCATATGGAGATGCTTACCATGGCTACT 360
Db 301 CCCCCTTTACAGCCCGAGCTGCCCGACAGACCCAGCATATGGAGATGCTTACCATGGCTACT 360
QY 361 GGCAGCAGGATATATATCTCTCTGAACGAAACTACGGCACTCGAGATGACTTTGAAGCGC 420
Db 361 GGCAGCAGGATATATATCTCTCTGAACGAAACTACGGCACTCGAGATGACTTTGAAGCGC 420
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGATCTTTATGTCGATGTTGGTGTAAACCA 480
Db 421 TCTCTTCGGCCCTTCATGAGAGGGGATGATCTTTATGTCGATGTTGGTGTAAACCA 480
QY 481 TGGGCTATGATGGAGCGGAGTACTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCC 540
Db 481 TGGGCTATGATGGAGCGGAGTACTCAGTCGATTAACAGTGTGTTTAAACCGTTTCAGTTCCC 540
QY 541 AAGACTACTTCAACCCGTTCTGTTTCAATTAAGATCAGATCAGGTTGAGG 600
Db 541 AAGACTACTTCAACCCGTTCTGTTTCAATTAAGATCAGATCAGGTTGAGG 600
QY 601 ATTGCTGGCTAGGAGTAACACTGTCTCTTGGCTGATCTCGATACCAACCAAGGATGTCG 660
Db 601 ATTGCTGGCTAGGAGTAACACTGTCTCTTGGCTGATCTCGATACCAACCAAGGATGTCG 660

```
QY 661 TCAGGAATGAATGGTACGACTGGTGGGATCATTTGTATCGAATCACTCAATTCAGCGCC 720
Db 661 TCAGGAATGAATGGTACGACTGGTGGGATCATTTGTATCGAATCACTCAATTCAGCGCC 720
QY 721 TCCGTATCGACACAGTAAACACGCTCCAGAAGGACTTCTGGCCCGGGTACAAACAAGCGC 780
Db 721 TCCGTATCGACACAGTAAACACGCTCCAGAAGGACTTCTGGCCCGGGTACAAACAAGCGC 780
QY 781 CAGGCGTGATCTGTATTCGGGAGGTGCTCGACGGTGATCCGGCTACACTTGTCCCTACC 840
Db 781 CAGGCGTGATCTGTATTCGGGAGGTGCTCGACGGTGATCCGGCTACACTTGTCCCTACC 840
QY 841 AGAACGTCATGGACGGGTACTGAACATATCCCAATTTACTATCACTCTCTCAAGCCCTCA 900
Db 841 AGAACGTCATGGACGGGTACTGAACATATCCCAATTTACTATCACTCTCTCAAGCCCTCA 900
QY 901 AGTCAACCTCCGCGACATGGACGACTCTCAACATGATCAACCGCTCAATTCGACT 960
Db 901 AGTCAACCTCCGCGACATGGACGACTCTCAACATGATCAACCGCTCAATTCGACT 960
QY 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAACCAACGAAACCCACGGTTCGTT 1020
Db 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAACCAACGAAACCCACGGTTCGTT 1020
QY 1021 CTTACACCAACGACATAGCCCTCGCAAGAACGTCGAGGATTCATCTCCTCAACGAC 1080
Db 1021 CTTACACCAACGACATAGCCCTCGCAAGAACGTCGAGGATTCATCTCCTCAACGAC 1080
QY 1081 GAATCCCATCATCTACGCGGCGCAAGAACGACACTACGCGCGGAAACGACCCCGCA 1140
Db 1081 GAATCCCATCATCTACGCGGCGCAAGAACGACACTACGCGCGGAAACGACCCCGCA 1140
QY 1141 ACCGGAAGCAACCTGGCTCTCGGGTACCAGCAACGACGAGCTGTACAAAGTTAATTG 1200
Db 1141 ACCGGAAGCAACCTGGCTCTCGGGTACCAGCAACGACGAGCTGTACAAAGTTAATTG 1200
QY 1201 CCTCCGGAACGCAATCCGGAATATGCCATTTAGCAAGATACAGGATTCGTGACCTACA 1260
Db 1201 CCTCCGGAACGCAATCCGGAATATGCCATTTAGCAAGATACAGGATTCGTGACCTACA 1260
QY 1261 AGAATCGGCCCATCTACAAAGACGACACCAACGATCGCCATGCGCAAGGCGACAGATGGGT 1320
Db 1261 AGAATCGGCCCATCTACAAAGACGACACCAACGATCGCCATGCGCAAGGCGACAGATGGGT 1320
QY 1321 CGCAGATCGTGACTATCTTTGTCAACAAAGGTGCTTCGGGTGATTCGTATACCTCTCT 1380
Db 1321 CGCAGATCGTGACTATCTTTGTCAACAAAGGTGCTTCGGGTGATTCGTATACCTCTCT 1380
QY 1381 TGAGTGGTGGGTTACACAGCCGCGCAGCAATTTGACGGAGTTCATTTGGTGCACAGCG 1440
Db 1381 TGAGTGGTGGGTTACACAGCCGCGCAGCAATTTGACGGAGTTCATTTGGTGCACAGCG 1440
QY 1441 TGACGGTTCGTTCCGATGGAATGTCCCTGTTCTATGCGAGTGGGCTACCTAGGGTAT 1500
Db 1441 TGACGGTTCGTTCCGATGGAATGTCCCTGTTCTATGCGAGTGGGCTACCTAGGGTAT 1500
QY 1501 TGTATCCGACTGAGAGTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGA 1560
Db 1501 TGTATCCGACTGAGAGTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGA 1560
QY 1561 GTATATGATGGTACTGCTATTTCAATCTGGCATTTGGAACAGTGGATTTGATGATACA 1620
Db 1561 GTATATGATGGTACTGCTATTTCAATCTGGCATTTGGAACAGTGGATTTGATGATACA 1620
QY 1621 GTTGGAGTCGTTACTGCTGTCTATCCCTTATCTCTGATTTGTTTTCGAAACCTTAATG 1680
Db 1621 GTTGGAGTCGTTACTGCTGTCTATCCCTTATCTCTGATTTGTTTTCGAAACCTTAATG 1680
QY 1681 CCAAGCAACGCTAGTCTATTTATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
Db 1681 CCAAGCAACGCTAGTCTATTTATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
```

```
RESULT 2
US-09-710-339A-1
; Sequence 1, Application US/09710339A
; GENERAL INFORMATION:
; APPLICANT: Biesgard-Prantzen, Henrik
; APPLICANT: Pedersen, Sven
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Fungamyl-like Alpha-Amylase Variants
; FILE REFERENCE: 5835.200-US
; CURRENT APPLICATION NUMBER: US/09/710,339A
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/165,786
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Aspergillus Oryzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(1547)
; OTHER INFORMATION: mat_peptide
; NAME/KEY: mat_peptide
; LOCATION: (114)..()
; OTHER INFORMATION:
US-09-710-339A-1
```

```
Query Match 100.0%; Score 1734; DB 32; Length 1734;
Best Local Similarity 100.0%; Pred. No. 7.8e-252;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAAAGGCATTTATGATGTCG 60
Db 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAAAGGCATTTATGATGTCG 60
QY 61 CQTGTGGTCTCTATTTCTGTACGGCTTTCAAGTCGGGCACCTGCTTTGGCTGCAACGC 120
Db 61 CQTGTGGTCTCTATTTCTGTACGGCTTTCAAGTCGGGCACCTGCTTTGGCTGCAACGC 120
QY 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCTCTACGGATCGATTTGCAAGACGG 180
Db 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCTCTACGGATCGATTTGCAAGACGG 180
QY 181 ATGGGTTCGACGACTCGACTTCTTAATCTACGCGATCAGAAATACCTGTGTGAAACATGGC 240
Db 181 ATGGGTTCGACGACTCGACTTCTTAATCTACGCGATCAGAAATACCTGTGTGAAACATGGC 240
QY 241 AGGGCATCATCGAAGTTGGACTATATCCAGGGAATGGGCTTTCAGGCCATCTGGATCA 300
Db 241 AGGGCATCATCGAAGTTGGACTATATCCAGGGAATGGGCTTTCAGGCCATCTGGATCA 300
QY 301 CCCCCGTTACAGCCAGCTGCCCGACACCCAGCATATGAGATGCTTACCATGGCTACT 360
Db 301 CCCCCGTTACAGCCAGCTGCCCGACACCCAGCATATGAGATGCTTACCATGGCTACT 360
QY 361 GGCAGCAGGATATATCTCTCTGAACGAAACTACGCGACTCGAGATGACTTGAAGCGCG 420
Db 361 GGCAGCAGGATATATCTCTCTGAACGAAACTACGCGACTCGAGATGACTTGAAGCGCG 420
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGTGGTGTAAACATA 480
Db 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGTGGTGTAAACATA 480
QY 481 TGGGTATGATCGAGCGGTAGCTCAGTCGATACAGTGTGTTAAACCGTTCAGTTCCC 540
Db 481 TGGGTATGATCGAGCGGTAGCTCAGTCGATACAGTGTGTTAAACCGTTCAGTTCCC 540
QY 541 AAGACTACTTCCACCGCTTCTGTTTTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
Db 541 AAGACTACTTCCACCGCTTCTGTTTTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
```

601 ATTGCTGGCTAGGAGATAACACTGTCTCTGCTGATCTCGATACCAACCAAGGATGG 660
Db
601 ATTGCTGGCTAGGAGATAACACTGTCTCTGCTGATCTCGATACCAACCAAGGATGG 660
661 TCAAGAAATGAATGGTACGACTGGTGGGATCAATGGTATCGAACTACTCCATTTGACGGCC 720
Db
661 TCAAGAAATGAATGGTACGACTGGTGGGATCAATGGTATCGAACTACTCCATTTGACGGCC 720
721 TCCGTATCCACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGGTACAAACAAGCCG 780
Db
721 TCCGTATCCACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGGTACAAACAAGCCG 780
781 CAGGGCTGACTGTATCGCGGAGGCTCTCGACGGTGATCCGGCCTACACTTGTCCCTACC 840
Db
781 CAGGGCTGACTGTATCGCGGAGGCTCTCGACGGTGATCCGGCCTACACTTGTCCCTACC 840
841 AGAAGTATGAGACGGCGTACTGAACCTATCCCAATTTACTATCCCTCAACGGCTTCA 900
Db
841 AGAAGTATGAGACGGCGTACTGAACCTATCCCAATTTACTATCCCTCAACGGCTTCA 900
901 AGTCAACCTCCGGCAGCATGGAGCGACTCTCAACATGATCAACACCGTCAATCCGACT 960
Db
901 AGTCAACCTCCGGCAGCATGGAGCGACTCTCAACATGATCAACACCGTCAATCCGACT 960
961 GTCCAGACTCAACACTCTCTGGGCACTTCTCGAGAACCAACGACACCCACGGTTCGCTT 1020
Db
961 GTCCAGACTCAACACTCTCTGGGCACTTCTCGAGAACCAACGACACCCACGGTTCGCTT 1020
1021 CTTTACACCAAGCATAGCCCTCGCAAGAAAGTTCGACGATTCATCTCTCAACGACGG 1080
Db
1021 CTTTACACCAAGCATAGCCCTCGCAAGAAAGTTCGACGATTCATCTCTCAACGACGG 1080
1081 GAAATCCCATCATCTACCGCGGCAAGAACAGCACTACCGCGGGAACCGCCGCGA 1140
Db
1081 GAAATCCCATCATCTACCGCGGCAAGAACAGCACTACCGCGGGAACCGCCGCGA 1140
1141 ACCGGAAGCAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAAAGTTAAATG 1200
Db
1141 ACCGGAAGCAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAAAGTTAAATG 1200
1201 CTTCCGGAAGCAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAAAGTTAAATG 1260
Db
1201 CTTCCGGAAGCAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAAAGTTAAATG 1260
1261 AGAATCGGCCCATCTACAAAGCAAGCAACAGCTACCGACGAGGCAAGATGGGT 1320
Db
1261 AGAATCGGCCCATCTACAAAGCAAGCAACAGCTACCGACGAGGCAAGATGGGT 1320
1321 CGCAGATCGTGAATCTTGTCCAAAGGGTGTCTCGGGTGTATCGTATACCCCTCTCT 1380
Db
1321 CGCAGATCGTGAATCTTGTCCAAAGGGTGTCTCGGGTGTATCGTATACCCCTCTCT 1380
1381 TGAGTGGTGGGTTTACACAGCGGCGCAAGATTTGACGAGGTCTATGGCTGCAACGCG 1440
Db
1381 TGAGTGGTGGGTTTACACAGCGGCGCAAGATTTGACGAGGTCTATGGCTGCAACGCG 1440
1441 TGACGGTGGTTCGGATGGAATGTGCTTCTCTATGCGAGGTGGCTACCTAGGTTAT 1500
Db
1441 TGACGGTGGTTCGGATGGAATGTGCTTCTCTATGCGAGGTGGCTACCTAGGTTAT 1500
1501 TGTATCCGACTGAGAAATTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGGA 1560
Db
1501 TGTATCCGACTGAGAAATTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGGA 1560
1561 GTATATGATGGTACTGCTATTCATCTGGCAATTTGGAACAGTGTGAGTTGATGTACA 1620
Db
1561 GTATATGATGGTACTGCTATTCATCTGGCAATTTGGAACAGTGTGAGTTGATGTACA 1620
1621 GTTGGAGTGGTACTGCTGCTATCCCTTACTCTTCGATTTTTCGAAACCCCTAATG 1680
Db
1621 GTTGGAGTGGTACTGCTGCTATCCCTTACTCTTCGATTTTTCGAAACCCCTAATG 1680
1681 CCAAGCAGCTAGTCTTATTATAGGAAAAA 1734

Db 1681 CCAAGCAGCTAGTCTTATTATAGGAAAAA 1734
RESULT 3
US-09-710-339B-1
; Sequence 1, Application US/09710339B
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Prantzen, Henrik
; APPLICANT: Pedersen, Sven
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Fungamyl-like Alpha-Amylase Variants
; FILE REFERENCE: 5835.200-US
; CURRENT APPLICATION NUMBER: US/09/710,339B
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/165,786
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Aspergillus Oryzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(1547)
; OTHER INFORMATION: mat_peptide
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (114)..(1733)
US-09-710-339B-1
Query Match 100.0%; Score 1734; DB 32; Length 1734;
Best Local Similarity 100.0%; Pred. No. 7.8e-252; Indels 0; Gaps 0;
Matches 1734; Conservative 0; Mismatches 0;
QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCAATTTATGATGGTGG 60
Db 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCAATTTATGATGGTGG 60
QY 61 CTTGCTGGTCTCTATTTCTGTACGGCTTCAAGTTCGGGCACTCTGCTTTGGCTGCAACGC 120
Db 61 CTTGCTGGTCTCTATTTCTGTACGGCTTCAAGTTCGGGCACTCTGCTTTGGCTGCAACGC 120
QY 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCTCTACGGATCGATTTGCAAGACGG 180
Db 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCTCTACGGATCGATTTGCAAGACGG 180
QY 181 ATGGGTTCGACGACTCGGACTTGTAAATCTATCGCGATCAGAAATCTGTGTGGAAATGGC 240
Db 181 ATGGGTTCGACGACTCGGACTTGTAAATCTATCGCGATCAGAAATCTGTGTGGAAATGGC 240
QY 241 AGGGCATCATCGAAGTTGACTATATCCAGGGAATGGGCTTACAGGCATCTGGATCA 300
Db 241 AGGGCATCATCGAAGTTGACTATATCCAGGGAATGGGCTTACAGGCATCTGGATCA 300
QY 301 CCCCCGTTACAGCCAGCTGCCAGACACCGCATATGGAGATGCCTACCATGGCTACT 360
Db 301 CCCCCGTTACAGCCAGCTGCCAGACACCGCATATGGAGATGCCTACCATGGCTACT 360
QY 361 GGCACAGGATATATCTCTCTGAACGAAACTACGCGACTCGAGATGATCTTGAAGCGC 420
Db 361 GGCACAGGATATATCTCTCTGAACGAAACTACGCGACTCGAGATGATCTTGAAGCGC 420
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTTCGATGGTGTCTAACCATTA 480
Db 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTTCGATGGTGTCTAACCATTA 480
QY 481 TGGGCTATGATGGAGCGGCTAGCTCAGTCGATTAACAGTGTGTTAAACCGTTCAAGTTCCC 540
Db 481 TGGGCTATGATGGAGCGGCTAGCTCAGTCGATTAACAGTGTGTTAAACCGTTCAAGTTCCC 540
QY 541 AAGACTACTTCCACCCGTTCTGTTTTCATTTCAAACTATGAAGATCAGACTCAGGTTGAGG 600

Db 541 ||||| AGAGTACTTCCACCGGTTCTGTTTCAATGAAACTATGAAGATCAGACTCAGGTTGAGG 600
Qy 601 ATTGCTGGTAGGAGATAACACTGTCTCTTGGCTGTATCGATACCAACGAGGATGCG 660
Db 601 ATTGCTGGTAGGAGATAACACTGTCTCTTGGCTGTATCGATACCAACGAGGATGCG 660
Qy 661 TCAAGATGAATGGTACGACTGGGTGGGATCATTGGTATCGAACTACTCCATTGACGGCC 720
Db 661 TCAAGATGAATGGTACGACTGGGTGGGATCATTGGTATCGAACTACTCCATTGACGGCC 720
Qy 721 TCCGTATCGACACAGTAAACACGCTCCAGAGGACTTCTGCGCCGGGTACAAACAGCCG 780
Db 721 TCCGTATCGACACAGTAAACACGCTCCAGAGGACTTCTGCGCCGGGTACAAACAGCCG 780
Qy 781 CAGCGGTGTAATGATCGGGAGGTGCTCGACGGTGTATCGGGCTACACTTGTCCCTTACC 840
Db 781 CAGCGGTGTAATGATCGGGAGGTGCTCGACGGTGTATCGGGCTACACTTGTCCCTTACC 840
Qy 841 AGAAGTATGAGAGGGGTACTGAACTATCCATTTACTATCACTCTCAAGCCCTTCA 900
Db 841 AGAAGTATGAGAGGGGTACTGAACTATCCATTTACTATCACTCTCAAGCCCTTCA 900
Qy 901 AGTCAACCTCCGCGAGCATGAGACGCTCTACACATGATCAACACGTCAAATCCGACT 960
Db 901 AGTCAACCTCCGCGAGCATGAGACGCTCTACACATGATCAACACGTCAAATCCGACT 960
Qy 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAACACGACACCCACGGTTGCTT 1020
Db 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAACACGACACCCACGGTTGCTT 1020
Qy 1021 CTTACACCAACGATAGCCCTCGCAGAACGTCGCGAGGATTCATCCTCAAGACG 1080
Db 1021 CTTACACCAACGATAGCCCTCGCAGAACGTCGCGAGGATTCATCCTCAAGACG 1080
Qy 1081 GAATCCCATCATCTACGCGGCCAGAACAGCAGACTACGCGGGGAAAGCCCGCGA 1140
Db 1081 GAATCCCATCATCTACGCGGCCAGAACAGCAGACTACGCGGGGAAAGCCCGCGA 1140
Qy 1141 ACCGCGAAGCAACTCTCGGCTTACCCGACCGACGAGCTGTACAAGTTAATTG 1200
Db 1141 ACCGCGAAGCAACTCTCGGCTTACCCGACCGACGAGCTGTACAAGTTAATTG 1200
Qy 1201 CTTCCGCGAAGCAACTCTCGGCTTACCCGACCGACGAGCTGTACAAGTTAATTG 1260
Db 1201 CTTCCGCGAAGCAACTCTCGGCTTACCCGACCGACGAGCTGTACAAGTTAATTG 1260
Qy 1261 AGAAGTGGCCCATCTACAGAGAGACACAGGATCGCCATGCGGAGGACAGATGGGT 1320
Db 1261 AGAAGTGGCCCATCTACAGAGAGACACAGGATCGCCATGCGGAGGACAGATGGGT 1320
Qy 1321 CGCAGATCGTGACTATCTTGTCAACAAAGGCTGCTTGGGTGATTCGTATACCTCTCCT 1380
Db 1321 CGCAGATCGTGACTATCTTGTCAACAAAGGCTGCTTGGGTGATTCGTATACCTCTCCT 1380
Qy 1381 TGAGTGTGCGGTTACAGCGCGCCAGCAATTGACGAGGCTATTGGCTGACGACCG 1440
Db 1381 TGAGTGTGCGGTTACAGCGCGCCAGCAATTGACGAGGCTATTGGCTGACGACCG 1440
Qy 1441 TGACGGTTGGTTGCGATGGAATGTGCTGTTCTTATGGCAGGTGGCTACCTAGGGTAT 1500
Db 1441 TGACGGTTGGTTGCGATGGAATGTGCTGTTCTTATGGCAGGTGGCTACCTAGGGTAT 1500
Qy 1501 TGTATCCGACTGAGAGTGTGGCAGGTAGCAAGATCTGTAGTAGCTGTGAAGGGTGAGA 1560
Db 1501 TGTATCCGACTGAGAGTGTGGCAGGTAGCAAGATCTGTAGTAGCTGTGAAGGGTGAGA 1560
Qy 1561 GTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGTGAGTTTGTATGATGAC 1620
Db 1561 GTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGTGAGTTTGTATGATGAC 1620
Qy 1621 GTTGGAGTGGTTACTGCTGTCATCCCTTATCTCTCGATTGTTTTCGAAACCTTAATG 1680

Db 1621 GTTGGAGTGGTTACTGCTGTCATCCCTTATCTCTTCGATTGTTTTCGAAACCTTAATG 1680
Qy 1681 CCAAGCACGCTAGTCTATTATAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
Db 1681 CCAAGCACGCTAGTCTATTATAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
RESULT 4
US-10-820-200-1
; Sequence 1, Application US/10820200
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Prantzen, Henrik
; APPLICANT: Pedersen, Sven
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Fungamyl-like Alpha-Amylase Variants
; FILE REFERENCE: 5835.200-US
; CURRENT APPLICATION NUMBER: US/10/820,200
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US/09/710,339
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/165,786
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Aspergillus Oryzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(1547)
; OTHER INFORMATION: mat_peptide
; NAME/KEY: mat_peptide
; LOCATION: (114)..(1733)
US-10-820-200-1
Query Match 100.0%; Score 1734; DB 63; Length 1734;
Best Local Similarity 100.0%; Pred. No. 7.8e-252;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCACATCAGCTCTCCCTTCTCTGACCAATTAACCCACAGAGGCAATTTATGATGTCG 60
Db 1 TCACATCAGCTCTCCCTTCTCTGACCAATTAACCCACAGAGGCAATTTATGATGTCG 60
Qy 61 CGTGGTGGTCTCTATTCTGTACGGCTTTCAGGTGCGGCACTGCTTTGGCTGCAACGC 120
Db 61 CGTGGTGGTCTCTATTCTGTACGGCTTTCAGGTGCGGCACTGCTTTGGCTGCAACGC 120
Qy 121 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTCAGGATCGATTTGCAAGGACGG 180
Db 121 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTCAGGATCGATTTGCAAGGACGG 180
Qy 181 ATGGTTCGAGCTGCGACTTGTATTAATCTGCGGATCAGAAATCTGTTGGGACATGGC 240
Db 181 ATGGTTCGAGCTGCGACTTGTATTAATCTGCGGATCAGAAATCTGTTGGGACATGGC 240
Qy 241 AGGGCATCATCAGCAAGTTGGACTATATCCAGGGAATGGGCTTTCAGAGCATCTGATCA 300
Db 241 AGGGCATCATCAGCAAGTTGGACTATATCCAGGGAATGGGCTTTCAGAGCATCTGATCA 300
Qy 301 CCCCCGTTACAGCCAGCTGCCCCAGACCAACCGCATATGGAGATGCTACCATGGCTTACT 360
Db 301 CCCCCGTTACAGCCAGCTGCCCCAGACCAACCGCATATGGAGATGCTACCATGGCTTACT 360
Qy 361 GGCAGCAGATATATATCTCTGACGAAACTACCGGCTGCGAGATGCTTGAAGCGGC 420
Db 361 GGCAGCAGATATATATCTCTGACGAAACTACCGGCTGCGAGATGCTTGAAGCGGC 420
Qy 421 TCTCTTCGCGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGTTGGTTGCTAAACATA 480
Db 421 TCTCTTCGCGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGTTGGTTGCTAAACATA 480

481 TGGGCTATGATGGAGGGGTAGCTCAGTCGATTTACAGTGTGTTTAAACGGTTCAAGTTC 540
Db
481 TGGGCTATGATGGAGGGGTAGCTCAGTCGATTTACAGTGTGTTTAAACGGTTCAAGTTC 540
QY
541 AAGACTACTTCCACCCGTTCTGTTTCATTTCAAAAATATGAAGATCAGACTCAGGTTGAGG 600
Db
541 AAGACTACTTCCACCCGTTCTGTTTCATTTCAAAAATATGAAGATCAGACTCAGGTTGAGG 600
QY
601 ATTGCTGGCTAGGAGATACACTGCTCTCCCTGCTGATCTCGATACCCCAAGGATGGG 660
Db
601 ATTGCTGGCTAGGAGATACACTGCTCTCCCTGCTGATCTCGATACCCCAAGGATGGG 660
QY
661 TCAAGAAATGAATGGTACGACTGGGTGGGATCAATTTGGTATCGAACTACTCCATTGACGGCC 720
Db
661 TCAAGAAATGAATGGTACGACTGGGTGGGATCAATTTGGTATCGAACTACTCCATTGACGGCC 720
QY
721 TCGGTATCGACACAGTAAACACGCTCCAGAGACTTTCTGGCCCGGGTACAAACAAGCGG 780
Db
721 TCGGTATCGACACAGTAAACACGCTCCAGAGAGACTTTCTGGCCCGGGTACAAACAAGCGG 780
QY
781 CAGGCGTGTACTGTATCGGCGAGGTGCTCGACGGTGTATCGGCGCTACACTTGTCCCTACC 840
Db
781 CAGGCGTGTACTGTATCGGCGAGGTGCTCGACGGTGTATCGGCGCTACACTTGTCCCTACC 840
QY
841 AGAACTCATGGAACGGCTACTGAACATATCCCAATTTACTATCCACTCTCTCAACGGCTTCA 900
Db
841 AGAACTCATGGAACGGCTACTGAACATATCCCAATTTACTATCCACTCTCTCAACGGCTTCA 900
QY
901 AGTCAACCTTCGGGACGATGAGAGCTCTCAACATGATCAACACCGTCAAAATCGGACT 960
Db
901 AGTCAACCTTCGGGACGATGAGAGCTCTCAACATGATCAACACCGTCAAAATCGGACT 960
QY
961 GTCCAGACTCAACACTCTCCGACATTTCTGAGAACACCAACGACACCCACCGTTCGCTT 1020
Db
961 GTCCAGACTCAACACTCTCCGACATTTCTGAGAACACCAACGACACCCACCGTTCGCTT 1020
QY
1021 CTTTACACCAACGACATAGCCCTCGCCAAAGAGTTCGACGATTCATCTCTCAACGACG 1080
Db
1021 CTTTACACCAACGACATAGCCCTCGCCAAAGAGTTCGACGATTCATCTCTCAACGACG 1080
QY
1081 GAATCCCATCATCTTACGCGGCGCAAGAACAGACTACGCGGCGGAAACGACCCCGCA 1140
Db
1081 GAATCCCATCATCTTACGCGGCGCAAGAACAGACTACGCGGCGGAAACGACCCCGCA 1140
QY
1141 ACGCGAAGCAACCTGGCTCTCGGGTACCCGACCGACGAGCTGTACAAAGTTAAATG 1200
Db
1141 ACGCGAAGCAACCTGGCTCTCGGGTACCCGACCGACGAGCTGTACAAAGTTAAATG 1200
QY
1201 CCTCCGCGAACGCAATCCGGAACATATGCCATTTAGCAAAAGATACAGGATTCGTGACCTACA 1260
Db
1201 CCTCCGCGAACGCAATCCGGAACATATGCCATTTAGCAAAAGATACAGGATTCGTGACCTACA 1260
QY
1261 AGAATCGGCGCATCTTACAAAGACGACAAACGATTCGCCATTCGCCAAGGCGACAGATGGGT 1320
Db
1261 AGAATCGGCGCATCTTACAAAGACGACAAACGATTCGCCATTCGCCAAGGCGACAGATGGGT 1320
QY
1321 CGCAGATCTGACTATCTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCTCTCTCT 1380
Db
1321 CGCAGATCTGACTATCTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCTCTCTCT 1380
QY
1381 TGAGTGGTGGGGTTACACAGCGCGGCGAGCAATTTGACGGAGGTCAATTTGGCTGACGACCG 1440
Db
1381 TGAGTGGTGGGGTTACACAGCGCGGCGAGCAATTTGACGGAGGTCAATTTGGCTGACGACCG 1440
QY
1441 TGAACGTTGGTTGGGATGGAATATGTCCTGTTTCTATGAGGAGTGGGCTTACCTAGGGTAT 1500
Db
1441 TGAACGTTGGTTGGGATGGAATATGTCCTGTTTCTATGAGGAGTGGGCTTACCTAGGGTAT 1500
QY
1501 TGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGA 1560
Db
1501 TGTATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGA 1560
QY
1561 GTATATGATGGTACTGCTATTCAATCTGGCAATTTGGACAGTGAGTTTGAATGTATCA 1620

Db 1561 GTATATGATGGTACTGCTATTCAATCTGGCATTTGCACAGTCAGTTTGAATGTATCA 1620
QY 1621 GTTGGAGTCGTTACTGCTGTCATCCCTTATCTCTTATCTCTTCTGATTTGTTTTCGAACCCCTAATG 1680
Db 1621 GTTGGAGTCGTTACTGCTGTCATCCCTTATCTCTTCTGATTTGTTTTCGAACCCCTAATG 1680
QY 1681 CCAACACGCTAGTCTATTATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
Db 1681 CCAACACGCTAGTCTATTATAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1734
RESULT 5
US-09-533-559-5355
; Sequence 5355, Application US/09533559
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; EARLIER FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5355
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
US-09-533-559-5355
Query Match 98.4%; Score 1707; DB 24; Length 1914;
Best Local Similarity 100.0%; Pred. No. 9.1e-248;
Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAAAGGCAATTTATGATGGTGC 60
Db 3 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAAAGGCAATTTATGATGGTGC 62
QY 61 CGTGTGTGCTCTATTCTGTACGGCTTCAGTCCGCGACCTGCTTTGGCTGCAACGC 120
Db 63 CGTGTGTGCTCTATTCTGTACGGCTTCAGTCCGCGACCTGCTTTGGCTGCAACGC 122
QY 121 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTCACGGATCGATTTGCAAGGACGC 180
Db 123 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCTCACGGATCGATTTGCAAGGACGC 182
QY 181 ATGGGTGCAAGCTCGGACTTGTATACTACGCGATCAGAAATACCTGTGTGTGAAACATGGC 240
Db 183 ATGGGTGCAAGCTCGGACTTGTATACTACGCGATCAGAAATACCTGTGTGTGAAACATGGC 242
QY 241 AGGGCATCATCGAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCATCTGGATCA 300
Db 243 AGGGCATCATCGAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCATCTGGATCA 302
QY 301 CCCCCGTTACAGCCAGCTGCCCCAGACACACCGCATATGAGATGCTTACCATGGCTACT 360
Db 303 CCCCCGTTACAGCCAGCTGCCCCAGACACACCGCATATGAGATGCTTACCATGGCTACT 362
QY 361 GGCACAGAGATATATCTCTGTGAACGAAACTACGCGATCAGAGATGATGGAAGCGC 420
Db 363 GGCACAGAGATATATCTCTGTGAACGAAACTACGCGATCAGAGATGATGGAAGCGC 422
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGATCTTTATGTCGATGTTGTTGCTTAACCAT 480
Db 423 TCTCTTCGGCCCTTCATGAGAGGGGATGATCTTTATGTCGATGTTGTTGCTTAACCAT 482

```
QY 481 TGGGCTATGATGAGCGGTAGCTCAGTCGATTACAGTGTGTTTAAACGGTTCAAGTCCC 540
Db 483 TGGGCTATGATGAGCGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTCAAGTCCC 542
QY 541 AAGACTACTTCCACCCGTTCTGTTTCAATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
Db 543 AAGACTACTTCCACCCGTTCTGTTTCAATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 602
QY 601 ATTGCTGGCTAGGAGATAACAATGTTCTCTCTGCTGATCTCGATACCAACAGGATGTGG 660
Db 603 ATTGCTGGCTAGGAGATAACAATGTTCTCTCTGCTGATCTCGATACCAACAGGATGTGG 662
QY 661 TCAAGATGAATGGTACGACTGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGCC 720
Db 663 TCAAGATGAATGGTACGACTGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGCC 722
QY 721 TCCGTATCGACACAGTAAACACAGTCCAGAAAGACTTCTGGCCCGGTACAAACAAGCCG 780
Db 723 TCCGTATCGACACAGTAAACACAGTCCAGAAAGACTTCTGGCCCGGTACAAACAAGCCG 782
QY 781 CAGGCGTGTAATGTATCGGCGAGTGTCTGACCGGTGATCCGGCTACACTTGTCCCTACC 840
Db 783 CAGGCGTGTAATGTATCGGCGAGTGTCTGACCGGTGATCCGGCTACACTTGTCCCTACC 842
QY 841 AGAAGCTCATGGAGCGGTACTGAATATCCCAATTTACTATTCACATCTCTCAAGCCCTTCA 900
Db 843 AGAAGCTCATGGAGCGGTACTGAATATCCCAATTTACTATTCACATCTCTCAAGCCCTTCA 902
QY 901 AGTCAACCTCCGACGATGAGACGCTCTACAAACATGATCAACCGTCAAAATCCGACT 960
Db 903 AGTCAACCTCCGACGATGAGACGCTCTACAAACATGATCAACCGTCAAAATCCGACT 962
QY 961 GTCCAGACTCAACACTCTCTGGGCACATTCGTTCGAGAACCAACGACCCACGGTTGGCTT 1020
Db 963 GTCCAGACTCAACACTCTCTGGGCACATTCGTTCGAGAACCAACGACCCACGGTTGGCTT 1022
QY 1021 CTTTACACCAAGCATAGCCCTCGCCAAAGAACGTCGAGCAATTCATCTCAAGACG 1080
Db 1023 CTTTACACCAAGCATAGCCCTCGCCAAAGAACGTCGAGCAATTCATCTCAAGACG 1082
QY 1081 GAATCCCATCATCTACGCGGCAAGAACAGACACTACGCGCGGAAAGACCCCGCA 1140
Db 1083 GAATCCCATCATCTACGCGGCAAGAACAGACACTACGCGCGGAAAGACCCCGCA 1142
QY 1141 ACCCGGAAGCAACTGCTCTCGGGTACCCGACCGACAGAGCTGTACAAGTTAATTG 1200
Db 1143 ACCCGGAAGCAACTGCTCTCGGGTACCCGACCGACAGAGCTGTACAAGTTAATTG 1202
QY 1201 CTTCCGGGAACGCAATCCGGAACATATGCCATTTAGCAAGATACAGGATTCGTGACCTACA 1260
Db 1203 CTTCCGGGAACGCAATCCGGAACATATGCCATTTAGCAAGATACAGGATTCGTGACCTACA 1262
QY 1261 AGAATGGCCCATCTACAAGACGACACACAGATCCGCAATGCGCAGGGCAGAGTGGGT 1320
Db 1263 AGAATGGCCCATCTACAAGACGACACACAGATCCGCAATGCGCAGGGCAGAGTGGGT 1322
QY 1321 CGCAGATCGTGAATCTTGTCCAAACAAAGGCTGCTTCGGGTGATTCGTATACCTCTCCT 1380
Db 1323 CGCAGATCGTGAATCTTGTCCAAACAAAGGCTGCTTCGGGTGATTCGTATACCTCTCCT 1382
QY 1381 TGAGTGTGGGGTTACAGCGCGGCAGCAATGTCAGAGGTCAATTTGGTGCACGACCG 1440
Db 1383 TGAGTGTGGGGTTACAGCGCGGCAGCAATGTCAGAGGTCAATTTGGTGCACGACCG 1442
QY 1441 TGACGGTTGGTTCGGATGGAAATGTCCTGTTCTTATGGCAGGTGGCTACCTAGGGTAT 1500
Db 1443 TGACGGTTGGTTCGGATGGAAATGTCCTGTTCTTATGGCAGGTGGCTACCTAGGGTAT 1502
QY 1501 TGTATCCGACTGAGAGTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGAGA 1560
Db 1503 TGTATCCGACTGAGAGTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGAGA 1562
QY 1561 GTATATGATGGTACTGCTATTCAATCTGGCAATTTGGACAGTGTGTTGAGTTTGATGTACA 1620
```

```
Db 1563 GTATATGATGGTACTGCTATTCAATCTGGCAATTTGGACAGTGTGTTGAGTTTGATGTACA 1622
QY 1621 GTTGGAGTGGTACTGCTGTCATCCCTTATCTCTTTCGATTTTTCGAACCCCTAATG 1680
Db 1623 GTTGGAGTGGTACTGCTGTCATCCCTTATCTCTTTCGATTTTTCGAACCCCTAATG 1682
QY 1681 CCAAGCACGCTAGTCTATTATATAGGAAA 1707
Db 1683 CCAAGCACGCTAGTCTATTATATAGGAAA 1709

RESULT 6
US-10-653-047-5355
; Sequence 5355, Application US/10653047
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5355
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
; US-10-653-047-5355

Query Match 98.4%; Score 1707; DB 58; Length 1914;
Best Local Similarity 100.0%; Pred. No. 9,1e-248;
Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCAATTTATGATGTCG 60
Db 3 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCACAGAGGCAATTTATGATGTCG 62
QY 61 CGTGGTGGTCTCTATTCTGTACGGCTTCAGGTCGGGCACTGCTTTGGCTGCAACGC 120
Db 63 CGTGGTGGTCTCTATTCTGTACGGCTTCAGGTCGGGCACTGCTTTGGCTGCAACGC 122
QY 121 CTGCGGACTGGCGATCGCAATCCATTATTTCTCTCTCAGGATCGATTTGCAAGACGG 180
Db 123 CTGCGGACTGGCGATCGCAATCCATTATTTCTCTCTCAGGATCGATTTGCAAGACGG 182
QY 181 ATGGGTGCGACGCTCGACTTGTAAATACTGCGGATCAGAAATACTGTGGTGAACATGGC 240
Db 183 ATGGGTGCGACGCTCGACTTGTAAATACTGCGGATCAGAAATACTGTGGTGAACATGGC 242
QY 241 AGGGCATCATCACAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCCATCTGGATCA 300
Db 243 AGGGCATCATCACAAGTTGGACTATATCCAGGGAATGGGCTTCACAGCCATCTGGATCA 302
QY 301 CCCCCTTACAGCCAGCTGCCAGACCAACCGCATATGGAGATGCTACCATGGCTACT 360
Db 303 CCCCCTTACAGCCAGCTGCCAGACCAACCGCATATGGAGATGCTACCATGGCTACT 362
QY 361 GGCAGCAGGATATATACTCTCTGAACGAAAACTACGCGACTGCAGATGACTTGAAGCGC 420
Db 363 GGCAGCAGGATATATACTCTCTGAACGAAAACTACGCGACTGCAGATGACTTGAAGCGC 422
QY 421 TCTCTTGGGCCCTTCATGAGAGGGGATGTATCTTATGGTTCGATGTTGCTTAACCATTA 480
```


Db 423 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTTATGGTCGATGGTGTGCTAACCAATA 482
Qy 481 TGGGCTATGATGAGGGGTAGCTCAGTCGATGCTGTTTAAACGGTTCAGTTCC 540
Db 483 TGGGCTATGATGAGGGGTAGCTCAGTCGATGCTGTTTAAACGGTTCAGTTCC 542
Qy 541 AAGACTACTTCCACCGGTTCTGTTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG 600
Db 543 AAGACTACTTCCACCGGTTCTGTTTCATTCAAACCTATGAAGATCAGACTCAGGTTGAGG 602
Qy 601 ATTGCTGGCTAGGAGATAACATGCTCTCTTGGCTGATCTCGATACCAACAAAGGATGGG 660
Db 603 ATTGCTGGCTAGGAGATAACATGCTCTCTTGGCTGATCTCGATACCAACAAAGGATGGG 662
Qy 661 TCAAGAACTCAATGGTAGGCTGGTGGGATCATTTGTTATCGAACTACTCCATTGTCAGCGCC 720
Db 663 TCAAGAACTCAATGGTAGGCTGGTGGGATCATTTGTTATCGAACTACTCCATTGTCAGCGCC 722
Qy 721 TCCGTATCGACACAGTAAACACAGTCCAGAAGGACTTCTGGCCCGGGTACAAACAAGCCG 780
Db 723 TCCGTATCGACACAGTAAACACAGTCCAGAAGGACTTCTGGCCCGGGTACAAACAAGCCG 782
Qy 781 CAGGCGTGTACTGTATCGGGAGGTGCTCGACGGTGATCCGGCTACACTTGTCCCTACC 840
Db 783 CAGGCGTGTACTGTATCGGGAGGTGCTCGACGGTGATCCGGCTACACTTGTCCCTACC 842
Qy 841 AGAAGCTGATGACGGGCTACTGAACCTATCCCAATTTACTATCCACTCCCTCAACGCCCTCA 900
Db 843 AGAAGCTGATGACGGGCTACTGAACCTATCCCAATTTACTATCCACTCCCTCAACGCCCTCA 902
Qy 901 AGTCAACCTCCGGCAGCATGGACGACTCTTCAACATGATCAACACCGTCAAAATCCGACT 960
Db 903 AGTCAACCTCCGGCAGCATGGACGACTCTTCAACATGATCAACACCGTCAAAATCCGACT 962
Qy 961 GTCCAGACTCAACATCTCTGGGACATTCGTGTGAGAACCAAGACACCCAGGTTGCTTT 1020
Db 963 GTCCAGACTCAACATCTCTGGGACATTCGTGTGAGAACCAAGACACCCAGGTTGCTTT 1022
Qy 1021 CTTTACACCAACGACATAGCCCTCGCCAAAGAGTGTGACGATTCATCATCTCCCAACGAG 1080
Db 1023 CTTTACACCAACGACATAGCCCTCGCCAAAGAGTGTGACGATTCATCATCTCCCAACGAG 1082
Qy 1081 GAATCCCATCATCTACCGCGGCCAAGAACAGCAGCTACCGCGCGGAAACGCCCGGCA 1140
Db 1083 GAATCCCATCATCTACCGCGGCCAAGAACAGCAGCTACCGCGCGGAAACGCCCGGCA 1142
Qy 1141 ACCGGAAGAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAGTTAAATTG 1200
Db 1143 ACCGGAAGAACCTGGCTCTCGGGTACCCGACGACGAGCTGTACAGTTAAATTG 1202
Qy 1201 CTTCCGGAACGCAATCCGGAATATGCCATTAGCAAAAGATCAGGATTCGTGACCTACA 1260
Db 1203 CTTCCGGAACGCAATCCGGAATATGCCATTAGCAAAAGATCAGGATTCGTGACCTACA 1262
Qy 1261 AGAATCGGCCCATCTTCAAAGACGACAAACGATCGCCATGCGCAAGGGCAGAGATGGGT 1320
Db 1263 AGAATCGGCCCATCTCAAAGACGACAAACGATCGCCATGCGCAAGGGCAGAGATGGGT 1322
Qy 1321 CGCAGATCGTGACTATCTTGTCCAAAGGGTCTTCCGGTGTATCGTATACCTCTCTCT 1380
Db 1323 CGCAGATCGTGACTATCTTGTCCAAAGGGTCTTCCGGTGTATCGTATACCTCTCTCT 1382
Qy 1381 TGAGTGGTGGGGTTACAGCGCGGCCAGCAATTTGACGAGGTCAATGGCTGCACGACCG 1440
Db 1383 TGAGTGGTGGGGTTACAGCGCGGCCAGCAATTTGACGAGGTCAATGGCTGCACGACCG 1442
Qy 1441 TGACGGTGGTTCGGATGGAATGTCCCTGTTCTTATGGCAGGTGGGCTACCTAGGGTAT 1500
Db 1443 TGACGGTGGTTCGGATGGAATGTCCCTGTTCTTATGGCAGGTGGGCTACCTAGGGTAT 1502
Qy 1501 TGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGA 1560
Db 1503 TGTATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTGGAGA 1562

Qy 1561 GTATATGATGGTACTGCTATTCAATCTGGCAATTGGACAGTGAAGTTGAGTTGATGTACA 1620
Db 1563 GTATATGATGGTACTGCTATTCAATCTGGCAATTGGACAGTGAAGTTGAGTTGATGTACA 1622
Qy 1621 GTTGAGTGGTTACTGCTGTCTATCCCTTATATCTCTTCGATTTGTTTTCGAACCCCTAATG 1680
Db 1623 GTTGAGTGGTTACTGCTGTCTATCCCTTATATCTCTTCGATTTGTTTTCGAACCCCTAATG 1682
Qy 1681 CCAAGCAGCTAGTCTATATATAGGAAA 1707
Db 1683 CCAAGCAGCTAGTCTATATATAGGAAA 1709

RESULT 7
PCT-US04-20499-39
; Sequence 39, Application PC/TUS0420499
; GENERAL INFORMATION:
; APPLICANT: Udagawa, Hiroaki
; APPLICANT: Taira, Rikako
; APPLICANT: Tkagi, Shinobu
; APPLICANT: Allain, Eric
; APPLICANT: Hjort, Carsten
; APPLICANT: Vikso-Nielsen, Anders
; TITLE OF INVENTION: HYBRID ENZYMES FOR STARCH PROCESSING
; FILE REFERENCE: 10490.500-US
; CURRENT APPLICATION NUMBER: PCT/US04/20499
; CURRENT FILING DATE: 2004-06-25
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1827)
PCT-US04-20499-39

Query Match 86.4%; Score 1497.8; DB 2; Length 1827;
Best Local Similarity 99.9%; Pred. No. 3.2e-216;
Matches 1499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 54 ATGGTCGGTGGTGGTCTCTATTCTGTACGGCTTCAGGTCGCGCACCTGCTTTGGCT 113
Db 1 ATGGTCGGTGGTGGTCTCTATTCTGTACGGCTTCAGGTCGCGCACCTGCTTTGGCT 60
Qy 114 GCAACGCCCTGCGGACTGCGGATCGCAATCCATTATTTCTTCTCACGGATCGATTGCA 173
Db 61 GCAACGCCCTGCGGACTGCGGATCGCAATCCATTATTTCTTCTCACGGATCGATTGCA 120
Qy 174 AGGACGGATGGGTGCGACGCTCGGACTTGTATACTGCGGATCAGAAATCTGTGTGGA 233
Db 121 AGGACGGATGGGTGCGACGCTCGGACTTGTATACTGCGGATCAGAAATCTGTGTGGA 180
Qy 234 ACATGGCAGGGCATCATCGACAAGTTGGACTATATCCAGGGATGGGCTTCACAGCCATC 293
Db 181 ACATGGCAGGGCATCATCGACAAGTTGGACTATATCCAGGGATGGGCTTCACAGCCATC 240
Qy 294 TGGATCACCCCGTTTACAGCCAGCTGCCCGCAGACCCAGCATATGGAGATGCTTACCAT 353
Db 241 TGGATCACCCCGTTTACAGCCAGCTGCCCGCAGACCCAGCATATGGAGATGCTTACCAT 300
Qy 354 GGCTACTGCGCAGCAGGATATATCTCTGTGAACGAAAACTACGGCACTGCAGATGACTTG 413
Db 301 GGCTACTGCGCAGCAGGATATATCTCTGTGAACGAAAACTACGGCACTGCAGATGACTTG 360
Qy 414 AAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTCCATGTTGCT 473
Db 361 AAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTCCATGTTGCT 420

Db 541 GTTGAGGATTGCTGGCTAGGAGATACACTGTCTCTCCCTGATCTCGATACCAACAG 600
Qy 654 GATGTGGTCAAGAAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATT 713
Db 601 GATGTGGTCAAGAAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATT 660
Qy 714 GACGGCTCCGATCGACACAGTAAACACGTCACGAAGGACTTCTGGCCCGGGTACAAAC 773
Db 661 GACGGCTCCGATCGACACAGTAAACACGTCACGAAGGACTTCTGGCCCGGGTACAAAC 720
Qy 774 AAAGCCGAGGCGTGTACTGTATCGGCGAGGTGCTCGAGGTGATCGCGCTTACACTTGT 833
Db 721 AAAGCCGAGGCGTGTACTGTATCGGCGAGGTGCTCGAGGTGATCGCGCTTACACTTGT 780
Qy 834 CCTTACCAGAACTCATGACGCGTACTGAACTATCCATTACTATCCACTCTCTCAAC 893
Db 781 CCTTACCAGAACTCATGACGCGTACTGAACTATCCATTACTATCCACTCTCTCAAC 840
Qy 894 GCCTTCAAGTCAACTCTCGGCGAGCATGGACGACCTCTACAAATGATCAACACCGTCAA 953
Db 841 GCCTTCAAGTCAACTCTCGGCGAGCATGGACGACCTCTACAAATGATCAACACCGTCAA 900
Qy 954 TCGGACTGTCAGACTCAACACTCTGGGCGACATTTGGTGGAGAACACGACCAACCGG 1013
Db 901 TCGGACTGTCAGACTCAACACTCTGGGCGACATTTGGTGGAGAACACGACCAACCGG 960
Qy 1014 TTTCGCTTCTTACACAAACGACATAGCCCTCGGCAAGACGTGCGAGCATTCATCATCTCTC 1073
Db 961 TTTCGCTTCTTACACAAACGACATAGCCCTCGGCAAGACGTGCGAGCATTCATCATCTCTC 1020
Qy 1074 AAGCAGGAATCCCATCATCTACGCGGCCAAGAACAGCACTAGCGCGCGGAAACGAC 1133
Db 1021 AAGCAGGAATCCCATCATCTAGCGGCCAAGAACAGCACTAGCGCGCGGAAACGAC 1080
Qy 1134 CCGCGAAACCGGAGCAACTCTGGCTCTCGGGTACCCGACGACGAGCTGTACAG 1193
Db 1081 CCGCGAAACCGGAGCAACTCTGGCTCTCGGGTACCCGACGACGAGCTGTACAG 1140
Qy 1194 TTAATTGCTCTCGCGAAACGCAATCCGGAATATGCCATTAGCAAGATACAGGATTCTGTG 1253
Db 1141 TTAATTGCTCTCGCGAAACGCAATCCGGAATATGCCATTAGCAAGATACAGGATTCTGTG 1200
Qy 1254 ACCTCAAGAACTGGCCCATCTTCAAAAGACGACAAACGATCGCCATGCGCAAGGGCACA 1313
Db 1201 ACCTCAAGAACTGGCCCATCTTCAAAAGACGACAAACGATCGCCATGCGCAAGGGCACA 1260
Qy 1314 GATGGTCCGACATCGTGAATCTATCTGTCCAAAGGGTGTCTCGGGTGAATTCGTATACC 1373
Db 1261 GATGGTCCGACATCGTGAATCTATCTGTCCAAAGGGTGTCTCGGGTGAATTCGTATACC 1320
Qy 1374 CTCTCTTGGTGGTGGGTTTACACAGCGCGCCAGCAATTCAGCGAGGTCAATTGGCTGC 1433
Db 1321 CTCTCTTGGTGGTGGGTTTACACAGCGCGCCAGCAATTCAGCGAGGTCAATTGGCTGC 1380
Qy 1434 ACGACCGTGACGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1493
Db 1381 ACGACCGTGACGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1440
Qy 1494 AGGGTATTGTATCCGACTGAGAAAGTGGGAGGTGGGAGGTGGGAGGTGGGAGGTGGGAGG 1553
Db 1441 AGGGTATTGTATCCGACTGAGAAAGTGGGAGGTGGGAGGTGGGAGGTGGGAGGTGGGAGG 1500
Qy 1554 G 1554
Db 1501 G 1501

RESULT 9

PCT-US04-20499-35
; Sequence 35, Application PC/TUS0420499
; GENERAL INFORMATION:
; APPLICANT: Udagawa, Hiroaki
; APPLICANT: Taira, Rikako

; APPLICANT: Tkagi, Shinobu
; APPLICANT: Allain, Eric
; APPLICANT: Hjort, Carsten
; APPLICANT: Vikso-Nielsen, Anders
; TITLE OF INVENTION: HYBRID ENZYMES FOR STARCH PROCESSING
; FILE REFERENCE: 10490.500-US
; CURRENT APPLICATION NUMBER: PCT/US04/20499
; CURRENT FILING DATE: 2004-06-25
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
PCT-US04-20499-35

Query Match 86.4%; Score 1497.4; DB 2; Length 1923;

Best Local Similarity 99.9%; Pred. No. 3.6e-216;
Matches 1498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 54 ATGGTCGGGTGGTGGTCTCTATTCTGTACGGCCCTTCAGGTCCGGCACCTCTCTTTGGCT 113
Db 1 ATGGTCGGGTGGTGGTCTCTATTCTGTACGGCCCTTCAGGTCCGGCACCTCTCTTTGGCT 60
Qy 114 GCAACGCTCGGGACTGGGATCGCAATCCATTATTTCTCTACGGATCGATTGCA 173
Db 61 GCAACGCTCGGGACTGGGATCGCAATCCATTATTTCTCTACGGATCGATTGCA 120
Qy 174 AGGACGGATGGTCCGACGACTGCGACTTGTATATCTGGGATCAGAAATCTGTGGTGA 233
Db 121 AGGACGGATGGTCCGACGACTGCGACTTGTATATCTGGGATCAGAAATCTGTGGTGA 180
Qy 234 ACATGGCAGGGCATCATCGACAAGTTGGACTATATCCAGGGAAATGGGCTTCACAGCCATC 293
Db 181 ACATGGCAGGGCATCATCGACAAGTTGGACTATATCCAGGGAAATGGGCTTCACAGCCATC 240
Qy 294 TGGATCACCCCGTTACAGCCCGAGCTGCCCCAGACACCGCATATGGAGATGCTTACCAT 353
Db 241 TGGATCACCCCGTTACAGCCCGAGCTGCCCCAGACACCGCATATGGAGATGCTTACCAT 300
Qy 354 GGTACTGCGACGAGATATATCTCTGAAACGAAACTACGGCACTGCAGATGACTTG 413
Db 301 GGTACTGCGACGAGATATATCTCTGAAACGAAACTACGGCACTGCAGATGACTTG 360
Qy 414 AAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTTATGTCGATGTGGTGTCT 473
Db 361 AAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTTATGTCGATGTGGTGTCT 420
Qy 474 AACCATATGGGTATATGAGCGGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTC 533
Db 421 AACCATATGGGTATATGAGCGGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTC 480
Qy 534 AGTTCCTCCAGACTCTTCACCCGTTCTGTTTCATTCAAAATATGAAAGATCAGACTCAG 593
Db 481 AGTTCCTCCAGACTCTTCACCCGTTCTGTTTCATTCAAAATATGAAAGATCAGACTCAG 540
Qy 594 GTTGAGGATTTGTCGTAGGAGATAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 653
Db 541 GTTGAGGATTTGTCGTAGGAGATAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Qy 654 GATGTGGTCAAGAAATGAATGGTATCGAATGGGTGGGATCATTTGGTATCGAACTACTCCATT 713
Db 601 GATGTGGTCAAGAAATGAATGGTATCGAATGGGTGGGATCATTTGGTATCGAACTACTCCATT 660
Qy 714 GACGGCTCCGATCGACACAGTAAACACGTCACGAAGGACTTCTGGCCCGGGTACAAAC 773
Db 661 GACGGCTCCGATCGACACAGTAAACACGTCACGAAGGACTTCTGGCCCGGGTACAAAC 720

```
QY 774 AAAGCCGAGCGGTGTACTGTATCGCGAGGTGCTCGACGGTGTATCGGCGCTACACTTGT 833
Db 721 AAAGCCGAGCGGTGTACTGTATCGCGAGGTGCTCGACGGTGTATCGGCGCTACACTTGT 780
QY 834 CCCTACCAAGACGTCTATGACGCGGTACTGAACTATCCCATTTACTATCCACTCCTCAAC 893
Db 781 CCCTACCAAGACGTCTATGACGCGGTACTGAACTATCCCATTTACTATCCACTCCTCAAC 840
QY 894 GCCTTAAGTCAACCTCCGCGAGCATGAGACGCTCTTACAAATGATCAACACCGTCAAA 953
Db 841 GCCTTAAGTCAACCTCCGCGAGCATGAGACGCTCTTACAAATGATCAACACCGTCAAA 900
QY 954 TCCGACTGTCCAGACTCAACACTCCTGGGACACATTCGTGAGAACACGACAAACCCACGG 1013
Db 901 TCCGACTGTCCAGACTCAACACTCCTGGGACACATTCGTGAGAACACGACAAACCCACGG 960
QY 1014 TTGCGCTTCTTACACCAACGACATAGCCCTCGCCAAAGAACTGCGAGCAATTCATCATCCTC 1073
Db 961 TTGCGCTTCTTACACCAACGACATAGCCCTCGCCAAAGAACTGCGAGCAATTCATCATCCTC 1020
QY 1074 AAGCAGGAATCCCATTCATCTACGCGCGCAAGAAACAGCACTACGCGCGCGGAACGAC 1133
Db 1021 AAGCAGGAATCCCATTCATCTACGCGCGCAAGAAACAGCACTACGCGCGCGGAACGAC 1080
QY 1134 CCGCGCAACCGCAAGCAACCTGGCTCTCGGGCTACCGACGACAGCGAGCTGTACAG 1193
Db 1081 CCGCGCAACCGCAAGCAACCTGGCTCTCGGGCTACCGACGACAGCGAGCTGTACAG 1140
QY 1194 TTAAATGGCTCCCGCAAGCAATCCGGAATATGCGCAATAGCAAGATACAGGATTCGTG 1253
Db 1141 TTAAATGGCTCCCGCAAGCAATCCGGAATATGCGCAATAGCAAGATACAGGATTCGTG 1200
QY 1254 ACCTCAAGAACTGGGCCATCTCAAAAGACACAAACGATGCCATGCCATGGCAAGGCACA 1313
Db 1201 ACCTCAAGAACTGGGCCATCTCAAAAGACACAAACGATGCCATGCCATGGCAAGGCACA 1260
QY 1314 GATGGCTCGAGATCGTACTATCTTGTCCAAAGGGTCTTCGGGTGATTCGTATACC 1373
Db 1261 GATGGCTCGAGATCGTACTATCTTGTCCAAAGGGTCTTCGGGTGATTCGTATACC 1320
QY 1374 CTCTCTTGAAGTGGTGGGTATACACGCGCGGACAGCAATGACGAGGTCATTGGCTGC 1433
Db 1321 CTCTCTTGAAGTGGTGGGTATACACGCGCGGACAGCAATGACGAGGTCATTGGCTGC 1380
QY 1434 ACACCGTGACGTTGGTTCGGAATGAAATGTGCTTCTTATGCGAGGTGGCTACCT 1493
Db 1381 ACACCGTGACGTTGGTTCGGAATGAAATGTGCTTCTTATGCGAGGTGGCTACCT 1440
QY 1494 AGGTTATGTATCCGACTGAGAGTGGCAGGTAGCAAGATCTGTAGTAGCTGTGAAG 1552
Db 1441 AGGTTATGTATCCGACTGAGAGTGGCAGGTAGCAAGATCTGTAGTAGCTGTGAAG 1499
```

RESULT 10

US-10-877-849-35

; Sequence 35, Application US/10877849

; GENERAL INFORMATION:

; APPLICANT: Udagawa, Hiroaki

; APPLICANT: Taira, Rikako

; APPLICANT: Taki, Shinobu

; APPLICANT: Allain, Eric

; APPLICANT: Hjord, Carsten

; APPLICANT: Vikso-Nielsen, Anders

; TITLE OF INVENTION: HYBRID ENZYMES FOR STARCH PROCESSING

; FILE REFERENCE: 10490.500-US

; CURRENT APPLICATION NUMBER: US/10/877, 849

; CURRENT FILING DATE: 2004-06-25

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 35

; LENGTH: 1923

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:
; OTHER INFORMATION: Artificial

; NAME/KEY: CDS

; LOCATION: (1)..(1923)

US-10-877-849-35

Query Match 86.4%; Score 1497.4; DB 64; Length 1923;

Best Local Similarity 99.9%; Pred. No. 3.6e-216;

Matches 1498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 54 ATGGTCGCGTGGTGGTCTCTATTTCGTACGGGCTTCAGGTGCGGCGACCTGCTTTGGCT 113
Db 1 ATGGTCGCGTGGTGGTCTCTATTTCGTACGGGCTTCAGGTGCGGCGACCTGCTTTGGCT 60
QY 114 GCAACGCTTCGCGACTGCGGATCGCAATCCATTTATTTCTTCTCAGGATCGATTGCA 173
Db 61 GCAACGCTTCGCGACTGCGGATCGCAATCCATTTATTTCTTCTCAGGATCGATTGCA 120
QY 174 AGGACGGATGGTTCGAGACTGCGGACTTGTATATCTGCGGATCAGAAATCTGTGTGA 233
Db 121 AGGACGGATGGTTCGAGACTGCGGACTTGTATATCTGCGGATCAGAAATCTGTGTGA 180
QY 234 ACATGCGAGGATCATCGCAAGTTGGACTATATCCAGGAAATGGCTTCACAGCCATC 293
Db 181 ACATGCGAGGATCATCGCAAGTTGGACTATATCCAGGAAATGGCTTCACAGCCATC 240
QY 294 TGGATCACTCCCGTTACAGCCGAGCTGCCCGAGACCCAGCATATGGAGATGCTTACCAT 353
Db 241 TGGATCACTCCCGTTACAGCCGAGCTGCCCGAGACCCAGCATATGGAGATGCTTACCAT 300
QY 354 GGTACTGCGAGGATATATCTCTCTGAAAGAAATACGCGCATCGAGATGAGATTG 413
Db 301 GGTACTGCGAGGATATATCTCTCTGAAAGAAATACGCGCATCGAGATGAGATTG 360
QY 414 AAGGCGCTCTTTCGCGCTTCATGAGAGGGGATGTATCTTATGTCGATGTTGGT 473
Db 361 AAGGCGCTCTTTCGCGCTTCATGAGAGGGGATGTATCTTATGTCGATGTTGGT 420
QY 474 AACCATATGGGCTATGATGGAGCGGTAGCTCAGTCGATTTACAGTGTGTTAAACCGTTC 533
Db 421 AACCATATGGGCTATGATGGAGCGGTAGCTCAGTCGATTTACAGTGTGTTAAACCGTTC 480
QY 534 AGTTCCCAAGACTACTTCCACCGGTTCTGTTTCATTTCAAAATATGAAGATCAGACTCAG 593
Db 481 AGTTCCCAAGACTACTTCCACCGGTTCTGTTTCATTTCAAAATATGAAGATCAGACTCAG 540
QY 594 GTTGAGGATTTGCTGGCTAGGATACACTGTCTCTGCTGCTGATCTCGATACCAAG 653
Db 541 GTTGAGGATTTGCTGGCTAGGATACACTGTCTCTGCTGCTGATCTCGATACCAAG 600
QY 654 GATGTGGTCAAGATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATT 713
Db 601 GATGTGGTCAAGATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATT 660
QY 714 GACGGCTCCGATTCGACACAGTAAACAGTCCAGAGGACTTCTGGCCCGGATCAAC 773
Db 661 GACGGCTCCGATTCGACACAGTAAACAGTCCAGAGGACTTCTGGCCCGGATCAAC 720
QY 774 AAAGCGGAGGCGGTGTACTGATCGCGAGGTGCTCGAGGGTGTATCCGGCTACACTTGT 833
Db 721 AAAGCGGAGGCGGTGTACTGATCGCGAGGTGCTCGAGGGTGTATCCGGCTACACTTGT 780
QY 834 CCCTACCAAGACGTCTATGAGCGCGGTACTGAACTATCCCATTTACTATCCACTCCTCAAC 893
Db 781 CCCTACCAAGACGTCTATGAGCGCGGTACTGAACTATCCCATTTACTATCCACTCCTCAAC 840
QY 894 GCCTTAAGTCAACCTCCGCGAGCATGAGACGCTCTTACAAATGATCAACACCGTCAAA 953
Db 841 GCCTTAAGTCAACCTCCGCGAGCATGAGACGCTCTTACAAATGATCAACACCGTCAAA 900
QY 954 TCCGACTGTCCAGACTCAACACTCCTGGGACACATTCGTGAGAACACGACAAACCCACGG 1013
```

```
Db 901 TCCGACTGTCAGACTCAACACTCTCTGGGACATTCGTGAGAACCAACGACAAACCCACGG 960
Qy 1014 TTCGCTTCTTACACAAACGACATAGCCCTCGCCAAAGACGTCGACGATTCATCATCTCTC 1073
Db 961 TTCGCTTCTTACACAAACGACATAGCCCTCGCCAAAGACGTCGACGATTCATCATCTCTC 1020
Qy 1074 AACGACGGAATCCCATCATCTACGCGGCCCAAGAACAGCACTACGCGCGGGAACGAC 1133
Db 1021 AACGACGGAATCCCATCATCTACGCGGCCCAAGAACAGCACTACGCGCGGGAACGAC 1080
Qy 1134 CCGCGGAACCGGAAGCAACCTGGCTCTCGGGCTTACCGACGACGAGCTGTACAG 1193
Db 1081 CCGCGGAACCGGAAGCAACCTGGCTCTCGGGCTTACCGACGACGAGCTGTACAG 1140
Qy 1194 TTAATTTGCTCCGCGGAACCAATCCGGAATATCGCACTATGCCATTTAGCAAGATACAGATTCTG 1253
Db 1141 TTAATTTGCTCCGCGGAACCAATCCGGAATATCGCACTATGCCATTTAGCAAGATACAGATTCTG 1200
Qy 1254 ACCTACAAGAACTGGGCCCATCTACAAGACGACACCAACGATGCCATGCGCAAGGGCACA 1313
Db 1201 ACCTACAAGAACTGGGCCCATCTACAAGACGACACCAACGATGCCATGCGCAAGGGCACA 1260
Qy 1314 GATGGTCCAGATCGTGAATCTTGTCCAAACAGGGTGTCTCGGGTGAATTCGTATACC 1373
Db 1261 GATGGTCCAGATCGTGAATCTTGTCCAAACAGGGTGTCTCGGGTGAATTCGTATACC 1320
Qy 1374 CTCTCTTGTGCTGGGTTACACAGCGGCCAGCAATTCACGAGGTCAATTCGCTGC 1433
Db 1321 CTCTCTTGTGCTGGGTTACACAGCGGCCAGCAATTCACGAGGTCAATTCGCTGC 1380
Qy 1434 ACACCGCTGACGGTGGTTCGGATGGAAATGTGCTTGTCTTATGGCAGGTGGCTACCT 1493
Db 1381 ACACCGTGAAGTGGTTCGGATGGAAATGTGCTTGTCTTATGGCAGGTGGCTACCT 1440
Qy 1494 AGGTATTGTATCCGACTGAGAAAGTTGGAGTAGCAAGATCTGTAGTAGCTCGTGAAG 1552
Db 1441 AGGTATTGTATCCGACTGAGAAAGTTGGAGTAGCAAGATCTGTAGTAGCTCGTGAAG 1499

RESULT 11
PCT-US02-27129-46
; Sequence 46, Application PC/TUS0227129
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: PCT/US02/27129
; CURRENT FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3285
; TYPE: DNA
; ORGANISM: Aspergillus shiroueami
PCT-US02-27129-46

Query Match 58.3%; Score 1011.2; DB 1; Length 3285;
Best Local Similarity 81.9%; Pred. No. 7.1e-143;
Matches 1166; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

Qy 114 GCAACGCTTCGGACTGGCGATCGCAATCCATTTATTTCTCTCACGGATCGATTGCA 173
Db 1 GCCACCCCGGCGACTGGCGCTCCAGTCCATCTACTTCTCTCACCGACCGCTTCGCC 60
Qy 174 AGGACGATGGTTCAGACGACTGGACTTGTATATCTGCGGATCAGAAATCTGTGTGGA 233
Db 61 CGCACCGAGGCTTCACACCGCCCACTGCAACACCGCGGACGAGAACTGTGCGCGGC 120
Qy 234 ACATGGCAGGATCATTCAGCAAGTTGGACTATATTCAGGGAATGGCTTCACGCCATC 293
Db 121 ACTGGCAGGATCATTCAGCAAGTTGGACTATATTCAGGGAATGGCTTCACGCCATC 180
Qy 294 TGGATCACCCCGCTTACAGCCCGAGCTGCCCGACACCGCATATGGAGATGCCTACCAT 353
```

```
Db 181 TGGATCACCCCGGTGACCGCCAGCTCCCGCAGACCAACCGCTACGGGACGCTACCAAC 240
Qy 354 GCCTACTCGCAGCAGATATATCTCTCTGACGAAACAACTACGGCAGCTGACAGTGTG 413
Db 241 GCCTACTCGCAGCAGATATCTCTCTCTGACGAAACAACTACGGCAGCTGACAGTGTG 300
Qy 414 AAGGCGCTCTCTTCCGCGCTTTCATGAGAGGGGATGTATCTTTATGTCATGTGTGTC 473
Db 301 AAGGCGCTCTCTTCCGCGCTTTCATGAGAGGGGATGTATCTTTATGTCATGTGTGTC 360
Qy 474 AACCATATGGCTATGATGAGAGGGGATGCTCAGTCGATTAACAGTGTGTAAACCGTTC 533
Db 361 AACCATATGGCTATGAGAGGGGCTCTCCGTCGACTACTCCGTCGTCGTCGTCGTCGTC 420
Qy 534 AGTTCCCAAGACTACTTCCACCGCTTCTGTTTTCATTAACAACTATGAGATCAGACTCAG 593
Db 421 TCCTCCCAAGACTACTTCCACCGCTTCTGTTTTCATTAACAACTATGAGATCAGACTCAG 480
Qy 594 GTTGAAGATTGCTGGCTAGGAGATTAACACTGTCTCTTTCCTTGCCTGATCTCGATACCAAC 653
Db 481 GTTGAAGATTGCTGGCTAGGAGATTAACACTGTCTCTTTCCTTGCCTGATCTCGATACCAAC 540
Qy 654 GATGTGGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 713
Db 541 GATGTGGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 714 GATGTGGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
Db 601 GATGTGGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 774 AAGGCGCGAGGCTGTACTGTATCCGCGAGGTGCTCGACGGTGTCTCGGCTTACACTTGT 833
Db 661 AAGGCGCGAGGCTGTACTGTATCCGCGAGGTGCTCGACGGTGTCTCGGCTTACACTTGT 720
Qy 834 CCCTACCAAGAGCTCATGGAAGGCTGATGAACTATCCATTTATTCATTCATTCATTCAT 893
Db 721 CCCTACCAAGAGCTCATGGAAGGCTGATGAACTATCCATTTATTCATTCATTCATTCAT 780
Qy 894 GCCTTCAAGTCAACCTCCGCGAGCTGAGGACCTTACAACTATGATCAACACCGCTCAAA 953
Db 781 GCCTTCAAGTCAACCTCCGCGAGCTGAGGACCTTACAACTATGATCAACACCGCTCAAA 840
Qy 954 TCCGACTGTCCAGACTCAACACTCTCTGGGCACTTCTGTCGAGAACCAACCAACCCACGG 1013
Db 841 TCCGACTGTCCGAGCTTCCACCTCTCTCGGCACTTCTGTCGAGAACCAACCAACCCACGG 900
Qy 1014 TTCGCTTCTTACCAACGACATAGCCCTTCGCAAGAGCTGCGAGCATTCATCATCTCTC 1073
Db 901 TTCGCTTCTTACCAACGACATAGCCCTTCGCAAGAGCTGCGAGCATTCATCATCTCTC 960
Qy 1074 AACGACGGAATCCCATCATCTACGCGGCCCAAGAACAGCACTACGCGCGGGAACGAC 1133
Db 961 AACGACGGAATCCCATCATCTACGCGGCCCAAGAACAGCACTACGCGCGGGAACGAC 1020
Qy 1134 CCGCGGAACCGGAAGCAACCTGGCTCTCGGGCTTACCGACGACGAGCTGTACAG 1193
Db 1021 CCGCGGAACCGGAAGCAACCTGGCTCTCGGGCTTACCGACGACGAGCTGTACAG 1080
Qy 1194 TTAATTTGCTCCGCGGAACCAATCCGGAATATCGCACTATGCCATTTAGCAAGATACAGATTCTG 1253
Db 1081 TTAATTTGCTCCGCGGAACCAATCCGGAATATCGCACTATGCCATTTAGCAAGATACAGATTCTG 1140
Qy 1254 ACCTACAAGAACTGGGCCCATCTACAAGACGACACCAACGATGCCATGCGCAAGGGCACA 1313
Db 1141 ACCTACAAGAACTGGGCCCATCTACAAGACGACACCAACGATGCCATGCGCAAGGGCACA 1200
Qy 1314 GATGGTCCAGATCGTGAATCTTGTCCAAACAGGGTGTCTCGGGTGAATTCGTATACC 1373
Db 1201 GATGGTCCAGATCGTGAATCTTGTCCAAACAGGGTGTCTCGGGTGAATTCGTATACC 1260
Qy 1374 CTCTCTTGTGCTGGGTTACACAGCGGCCAGCAATTCACGAGGTCAATTCGCTGC 1433
```

```
Db 1261 CTCTCCCTCTCGCGCGCGCTACACCGCGCGCAGAGCTACCGAGGTGATCGGCTGC 1320
Qy 1434 ACACCGGTGACCGGTTGTCGGATGGAATGTCCTTCTTATGGCAGGTGGCTACCT 1493
Db 1321 ACCACCGTACCGTGGGCTCCGACGGCAACATCGCGTGGCGGCTCCCG 1380
Qy 1494 AGGATTGATCCGACTGAGAGTTGGCAGGTAGCAAGATCTG 1537
Db 1381 CGGTGCTCTACCGACCGGAGAGCTCGCGGCTCCAGATATG 1424

RESULT 12
PCT-US04-07182A-46
; Sequence 46, Application PC/TUS0407182A
; GENERAL INFORMATION:
; APPLICANT: Lananhan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: PCT/US04/07182A
; CURRENT FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3285
; TYPE: DNA
; ORGANISM: Aspergillus shirousami
PCT-US04-07182A-46

Query Match 58.3%; Score 1011.2; DB 2; Length 3285;
Best Local Similarity 81.9%; Pred. No. 7.1e-143;
Matches 1166; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

Qy 114 GCAACGCTTCGGACTGGCGATCGCAATCCATTTATTTCTTCTCAGGATCGATTGCA 173
Db 1 GCCACCCCGCGACTGGCGCTCCAGTCCATCTACTTCTCTCACCAGCGCTTCGCG 60
Qy 174 AGGACGATGGGTGCGAGTGGACTTGTATATCTCGGATCAGAAATCTGTGTGGA 233
Db 61 CGACCGAGCGGTCCACCCCGCCACTGCAACACCGCGACAGAACTATCGCGCGGC 120
Qy 234 ACATGCGAGGCAATATCGACAGTTGGACTATATCCAGGGAATGGCTTCACAGCCATC 293
Db 121 ACTGCGAGGGATATATCGAAGCTCGACTATATCGAGGGATGGCTTCACCGGCATC 180
Qy 294 TGATCACCCTCGTTACGCCAGCTGCCCCAGACACCGCATATGGAGATGCTTACCAT 353
Db 181 TGATCACCCTCGTGACGCCAGCTCCCGCAGACACCGCATATCGGAGCGCTTACCAC 240
Qy 354 GGTACTGCGAGGATATATATCTCTTGAAGCAAACTACGGCACTCGAGATGACTTG 413
Db 241 GGTACTGCGAGGATATATCTCTTCAACGAGAACTACGGCACTCGGCACTCGGACCTC 300
Qy 414 AAGGCGCTCTTTCGGCCCTCATGAGAGGGGATGTATCTTATGTCGATGTTGCT 473
Db 301 AAGGCGCTCTTTCGGCCCTCATGAGAGGGGATGTATCTTATGTCGATGTTGCT 473
Qy 474 AACCATATGGGCTATGATGAGCGGTAGTCTAGTCCGATTTACAGTGTGTTTAAACCGTTC 533
Db 361 AACCATATGGGCTATGATGAGCGGTAGTCTAGTCCGATTTACAGTGTGTTTAAACCGTTC 420
Qy 534 AGTTCCCAAGACTATCTTCAACCGGTTCTGTTTCAATTCATAAATATGAAGATGACTGAG 593
Db 421 TCTTCCAGGACTATCTTCAACCGGTTCTGTTTCAATTCATAAATATGAAGATGACTGAG 480
Qy 594 GTTGAGGATTTGGCTAGGAGATACACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 653
Db 481 GTTGAGGATTTGGCTAGGAGATACACTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 654 GATGCTCAAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713
Db 541 GACGTGTGAGAACGAGTGGTACGACTGGTGGGCTCCCTCGTGTCCAACTACTTCCATC 600
Qy 714 GACGGCTCCGTATCGACACAGTAAACAGCTCCAGAGGACTTTCGGCGCGGGTCAAC 773
```

```
Db 601 GACGGCTCCGCTATCGACACCGTGAGACAGCTTTCGGCCGGGCTACAAC 660
Qy 774 AAAGCCGAGGCGGTGTACTGTATCGCGAGGTGCTCGAGGTGATCCGGCTTACACTGT 833
Db 661 AAGCCGCGCGCGGTGTACTGTATCGCGAGGTGCTCGAGGTGATCCGGCTTACACTGT 720
Qy 834 CCTTACCAAGAGTGTATGAGCGGCTGTAAGTATCCCATTTACTATTCACCTCTCAAC 893
Db 721 CGGTACCAAGAGTGTATGAGCGGCTGTAAGTATCCCATTTACTATTCACCTCTCAAC 780
Qy 894 GCCTTCAAGTCAACCTCCGCGAGCTGAGACAGCTTCAACATGATCAACCGCTCAA 953
Db 781 GCCTTCAAGTCCACCTCCGCGCTGATGAGACAGCTTCAACATGATCAACCGTGAAG 840
Qy 954 TCCGACTGTCCAGACTCAACACTCTCTGGGACATTCGTGGAGAACCAACCAACCGG 1013
Db 841 TCCGACTGTCCGAGCTCCACCTCTCTCGGACCTTCGTGGAGAACCAACCAACCGG 900
Qy 1014 TTGCTTCTTACACCAAGACATAGCCCTCGGCAAGAACGTGCGAGATTCATCATCTC 1073
Db 901 TTGCTTCTTACACCAAGACATAGCCCTCGGCAAGAACGTGCGGCTTCATCATCTC 960
Qy 1074 AACGAGGAATCCCATCATCTACGCGGCGCAGACAGCACTACCGCGGCAAGGAC 1133
Db 961 AACGAGGAATCCCATCATCTACGCGGCGCAGGAGCACTACCGCGGCGCAAGGAC 1020
Qy 1134 CCGCGAACCGCGAGCAACCTGGCTCTCGGGTACCCGACCGACAGCGAGCTGTACAAG 1193
Db 1021 CCGCGAACCGCGAGCGGCACTGGCTCTCGGCTACCCGACCGACTCCGAGCTGTACAAG 1080
Qy 1194 TTAATGCTTCCGCGAACCGCAATTCGGAATCTATGCTTATGCAAGATACAGGATTCGT 1253
Db 1081 CTCATCGCTTCGCGAACCGCAATTCGGAATCTATGCTTATGCAAGATACAGGATTCGT 1140
Qy 1254 ACCTACAGAACTGGCCCATCTACAGAGCAGACACCACTACCGCGGCAAGGACCA 1313
Db 1141 ACCTACAGAACTGGCCCATCTACAGAGCAGACACCACTACCGCGGCAAGGACCA 1200
Qy 1314 GATGGTCCAGATCGTGAATCTTGTCCAAAGGGTCTTCGGGTGATTCGTATACC 1373
Db 1201 GACGCTCCAGATCGTGAATCTTGTCCAAAGGGTCTTCGGGTGATTCGTATACC 1260
Qy 1374 CTCCTCTTGAATGCTGGGTTCACAGCGGCGGCAATTCAGGAGGTTCATTCGCTGC 1433
Db 1261 CTCCTCTTGAATGCTGGGTTCACAGCGGCGGCAATTCAGGAGGTTCATTCGCTGC 1320
Qy 1434 ACAGCGTCAAGTGGTTCGGATGGAATGTCCTGTTTCTTATGCGAGTGGCTACCT 1493
Db 1321 ACAGCGTCAAGTGGTTCGGATGGAATGTCCTGTTTCTTATGCGAGTGGCTACCT 1380
Qy 1494 AGGATTTGTATCCGACTGAGAAATGTCGAGGTGAGCAAGATCTG 1537
Db 1381 CGGTGCTCTACCGACCGAGAGCTCGCGGCTCCAGATATG 1424
```

```
RESULT 13
US-10-228-063-46
; Sequence 46, Application US/10228063
; GENERAL INFORMATION:
; APPLICANT: Lananhan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 3285
; TYPE: DNA
; ORGANISM: Aspergillus shirousami
US-10-228-063-46
```

Query Match		58.3%;	Score 1011.2;	DB 50;	Length 3285;
Best Local Similarity		81.9%;	Pred. No. 7.1e-143;		
Matches 1166;		Conservative 0;	Mismatches 258;	Indels 0;	Gaps 0;
QY	114	GCAACGCTGGCGACTGGCGATCGCAATCCATTTATTTCTCTCAGGATCGATTGCA	173		
DB	1	GCACCCCGCGGACTGGCGCTCCAGTCCATCTACTTCTCTCACCAGCGCTCGCC	60		
QY	174	AGSACGGATGGGTTCGACGACTGGACTTGTATATCTGCGGATCAGAAATACTTGTGTGA	233		
DB	61	CGACCGAGGCTTCAACACCGCCACTCGAACCCGCCGACAGAACTACTTGGCGGC	120		
QY	234	ACATGSCAGGATCATCGACAAGTTGGACTATATCCAGGGAATGGCTTTCACAGCAATC	293		
DB	121	ACCTGSCAGGATCATCGACAAGCTCGACTATATCCAGGATGGCTTTCACCGCAATC	180		
QY	294	TGGATCACCCCGTTACAGCCGAGTGGCCCGACACACCGCATATGGAGATGCTTACCAT	353		
DB	181	TGGATCACCCCGTTGACCGCCAGCTCCCGCAGACACACCGCTTACGGCGAGCTTACCAC	240		
QY	354	GGCTACTGSCAGGATATATCTCTGAAACGAAACTACGGCACTGCGACTGACTTGTG	413		
DB	241	GGCTACTGSCAGGATATCTCTCTCAAGGAACTACGGCACTGCGCACTGCGCACTTC	300		
QY	414	AAGGCGCTCTCTTGGCCCTTCATGAGAGGGGATGTATCTTATGTCGATGTGGTTGCT	473		
DB	301	AAGGCGCTCTCTTGGCCCTTCACGAGCGCGCATGTACTCTCATGGTGAAGTGGTGGC	360		
QY	474	AACATATGGGCTATGATGAGGGGTAGCTCAGTCGATTAAGTGTGTTTAAACGGTTC	533		
DB	361	AACCATGGGCTATGATGAGGGGTAGCTCAGTCGATTAAGTGTGTTTAAACGGTTC	420		
QY	534	AGTTCACAGACTACTTCCACCGCTCTGTTTCATTCACAACTATGAAGTACAGACTCAG	593		
DB	421	TGCTCCAGACTACTTCCACCGCTCTGTTTCATTCACAACTATGAAGTACAGACTCAG	480		
QY	594	GTTGAGGATGTGGTAGGAGATAACACTGTCTCTTCCCTGATCTCGATACCAACCAAG	653		
DB	481	GTGGAGACTGTGGCTCGGCGACACACCGTGTCTCTCCGAGCTCGACACCAACCAAG	540		
QY	654	GATGGGTCAAGATGAATGGTATGATCGATGGGTGGATTCATTTGATTCGAATCTCCATT	713		
DB	541	GACGTGGTGAAGAACGAGTGGTACGACTGGGTGGCTCTCTGTTCCAACTACTCCATC	600		
QY	714	GACGCTCCGTATCGACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGTTACAC	773		
DB	601	GACGCTCCGTATCGACACAGTAAACACGCTCCAGAGGACTTCTGGCCCGGTTACAC	660		
QY	774	AAAGCCGAGGCTGTACTGTATCGCGGAGGTGCTCGACGGTATCCGGCTTACACTTGT	833		
DB	661	AAAGCCGCGGCTGTACTGTATCGCGGAGGTGCTCGACGGTATCCGGCTTACACTTGT	720		
QY	834	CCCTACAGAGCTCATGAGCGGCTATGAACTATCCCATTTATCTATCTCCTCAAC	893		
DB	721	CGGTACAGAGCTCATGAGCGGCTGTCAACTACCCGATCTACTACCGCTCTCTCAAC	780		
QY	894	GCCTTCACTCACTCCGCGACGTGAGACGCTTCAACATGATCAACACCGTCAA	953		
DB	781	GCCTTCACTCACTCCGCGACGTGAGACGCTTCAACATGATCAACACCGTGAAG	840		
QY	954	TCCGACTGTCCAGACTCAACACTCTCTGGGCACATTCGTGAGAACACACACCAACCGG	1013		
DB	841	TCCGACTGTCCGAGCTTCCACCTCTCTCGGACCTTCTGTGAGAACACACCAACCGGCG	900		
QY	1014	TTGCGCTTCTTACACCAACGATAGCCCTTCGCCAAGAACGTGCGAGCATTCATCACTCTC	1073		
DB	901	TTGCGCTTCTTACACCAACGATAGCCCTTCGCCAAGAACGTGCGAGCATTCATCACTCTC	960		
QY	1074	AACGAGGAATCCCATCATCTACCGCCGCGCAGACACGCTACGCGCGCGGAACGAC	1133		
DB	961	AACGAGGAATCCCATCATCTACCGCCGCGCAGGAGACGCTACGCGCGCGCGACGAC	1020		
QY	1134	CCGCGAACCGGAGCAACCTGGCTCTCGGGCTACCCGACCGACAGCGAGCTGTCAAG	1193		

Db	1021	CGGGCAACCGGAGGCCACCTGGCTCTTCGGGCTACCCGACCGACTCCGAGCTGTACAAG	1080
QY	1194	TTAATTGCTCTCCGGAAACGCAATCCGAACTATGCCATTAGCAAAAGATACAGGATTCTGTG	1253
Db	1081	CTCATCGCTCTCCGCAACGCCATCGCRACTACGCCATCTCCAAGGACACCGGCTTCGTG	1140
QY	1254	ACCTACAAGAACTGGCCCATCTTACAAGACGACACAACGATCGCCATGCGCAAGGGGACA	1313
Db	1141	ACCTACAAGAACTGGCCGATCTACAAGGACGACACCAATCGCCATGCGCAAGGGGCACC	1200
QY	1314	GATGGGTGCGAGATCGTGACTATCTTGTCTCAACAGGGTGCTTCGGGTGATTCTGTATACC	1373
Db	1201	GACGGCTCCAGATCGTGACCATCTCTCCAACAGGGCGCCTCCGGGACTCCTTACACC	1260
QY	1374	CTCTCTTTGAGTGGTGGGGTTACACAGCGCGCCAGCAATTGACGGAGGTCAATTGGCTGC	1433
Db	1261	CTCTCTCCTCTCCGGCGCGGCTTACCGCGGCCAGCAGCTCACCGAGGTGATCGGCTGC	1320
QY	1434	ACGACCGTGACGGTTGGTTCGGATGGAATGTGCCTGTTCTTATGGCAGGTGGGCTACCT	1493
Db	1321	ACCACCGTGACCGTGGGCTCCGACGGCAACGTGCGGTGCGATGGCGGCGGCTCCCG	1380
QY	1494	AGGGTATTGTATCCGACTGAGAAATTGGCAGGTAGCAAGATCTG	1537
Db	1381	CGCGTGCTCTACCCGACCGAGAAGCTCGCGGCTCCAAGATATG	1424
RESULT 14			
PCT-US04-10126-17			
; Sequence 17, Application PC/TUS0410126			
; GENERAL INFORMATION:			
; APPLICANT: Novozymes Biotech, Inc.			
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient			
; FILE REFERENCE: 10345.204-WO			
; CURRENT APPLICATION NUMBER: PCT/US04/10126			
; CURRENT FILING DATE: 2004-04-06			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 17			
; LENGTH: 2443			
; TYPE: DNA			
; ORGANISM: Aspergillus niger			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (10)..(10)			
; OTHER INFORMATION: n=a,c,g or t			
PCT-US04-10126-17			

QY 246 ATCATCGACA----- 255
DB 451 ATCATCGACAAGGTAAATTGCCCTTTTATCAAAAAAAGAGAAAGAGAGAAAAA 510
QY 256 -----AGTTGGACTATATCCAGGGAATGGG 280
DB 511 TAAATAAAGAACTCTAGTCTTAACCATCAATAGTTGGACTATATCCAGGGAATGGG 570
QY 281 CTTACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCCAGACCAACCCATATGG 340
DB 571 CTTACAGCCATCTGGATCACCCCGTTACAGCCAGCTGCCCCAGACCAACCCATATGG 630
QY 341 AGATGCTTACCATGGCTACTGGCAGCAGGATAT----- 373
DB 631 AGATGCTTACCATGGCTACTGGCAGCAGGATATGTAAGTCGATTTCTTTAAATATCTACC 690
QY 374 -----ATACTCTCTGAACGAAA 391
DB 691 TGTCTATCTTTTACATCAATATGAACCTAACCTTGATGGTTTGTAGATACTCTCTGAACGAAA 750
QY 392 CTACGGCACTCGAGATGACTTGAAGGGCTCTCTTCGGCCCTTCATGAGAGGGGGATGTA 451
DB 751 CTACGGCACTCGAGATGACTTGAAGGGCTCTCTTCGGCCCTTCATGAGAGGGGGATGTA 810
QY 452 TCTTATGGTCGATGTGGTTGCTAACCATAT----- 481
DB 811 TCTTATGGTCGATGTGGTTGCTAACCATATGTTGTTGGTGTCTTTGCAACTGACTTCGGC 870
QY 482 -----GGCTATGATGAGCGGTAGC 503
DB 871 GATATGGTTCAATTTCAGTACTGCAATAGTAATATACAGGGCTATGATGAGCGGTAGC 930
QY 504 TCAGTCGATTAACAGTGTGTTAAACCGTTCCAGTTCCTCAAGACTACTTCCACCCGTTCTGT 563
DB 931 TCAGTCGATTAACAGTGTGTTAAACCGTTCCAGTTCCTCAAGACTACTTCCACCCGTTCTGT 990
QY 564 TTCATTCAAACTATGAAGATCAGACTCAGTTGAGGATTCCTGGCTAGAGATAACACT 623
DB 991 TTCATTCAAACTATGAAGATCAGACTCAGTTGAGGATTCCTGGCTAGAGATAACACT 1050
QY 624 GTCTCTTCGCTGATCTCGATACCAAGGATGTGTCAGAAATGATGATGATGATGATG 683
DB 1051 GTCTCTTCGCTGATCTCGATACCAAGGATGTGTCAGAAATGATGATGATGATGATG 1110
QY 684 GTGGGATCATTTGGTATCGAACTACTCCA----- 711
DB 1111 GTGGGATCATTTGGTATCGAACTACTCCAGTAAGATATTTCTCCCTCATTTCTCAACTGG 1170
QY 712 -----TTGACGGCTCCGTPATCGACACAGTAAACACAGT 745
DB 1171 CTGATCGATGATCTTACGAAATCAGTTGAGCGCTCCGTPATCGACACAGTAAACACAGT 1230
QY 746 CCAGAGGACTTCTGGCCCGGTACAAACAAAGCGCGAGGCTGTACTGTATCGCGAGGT 805
DB 1231 CCAGAGGACTTCTGGCCCGGTACAAACAAAGCGCGAGGCTGTACTGTATCGCGAGGT 1290
QY 806 GCTCGAGGATCGCGCTTACACTTGTCTCCCTACCAAGAGCTCATGACGCGGTACTGAA 865
DB 1291 GCTCGAGGATCGCGCTTACACTTGTCTCCCTACCAAGAGCTCATGACGCGGTACTGAA 1350
QY 866 CTATCCCAT----- 874
DB 1351 CTATCCCATGATGTTCTCCCAACCATGAGCTTCTTGAAGTCTCATCTCTCAACGAA 1410
QY 875 -----TTACTATCACTCTCTCAAGCGCTTCAAGTCAACCTCCGGCAGCATG 920
DB 1411 ACGGCTAAACCAAGTTACTATCACTCTCAAGCGCTTCAAGTCAACCTCCGGCAGCATG 1470
QY 921 GAGACCTCTACACATGATCAACCGTCAATCCGACTTCCAGACTCAACACTCTCTG 980
DB 1471 GAGACCTCTACACATGATCAACCGTCAATCCGACTTCCAGACTTCCAGACTCAACACTCTCTG 1530

QY 981 GGCACATTCTGTCGAGAACCAACCAACCCACCGTTCCGCTTC----- 1021
DB 1531 GGCACATTCTGTCGAGAACCAACCAACCCACCGTTCCGCTTCGTAAGTCTTCCTTTTATT 1590
QY 1022 -----TTACACCAACGACA 1035
DB 1591 TTCCGTTCCCAATTTCCACACAGAACCCACCTTAACAAGAGCAAGTTACACCAACGACA 1650
QY 1036 TAGCCCTTCGCAAGAACGTCGCGAGCATTCATCATCTCAACGACGGAATCCCATCATCT 1095
DB 1651 TAGCCCTTCGCAAGAACGTCGCGAGCATTCATCATCTCAACGACGGAATCCCATCATCT 1710
QY 1096 ACGCCGGCCAAAGAACAGCACTACGCGCGGGAACCCCGCGAACCCCGGAGACCACT 1155
DB 1711 ACGCCGGCCAAAGAACAGCACTACGCGCGGGAACCCCGCGAACCCCGGAGACCACT 1770
QY 1156 GGTCTTCGGGCTACCCGACGACGAGCTGTACAAGTTAAATTGCTTCGCGAACGCAA 1215
DB 1771 GGTCTTCGGGCTACCCGACGACGAGCTGTACAAGTTAAATTGCTTCGCGAACGCAA 1830
QY 1216 TCCGGAATATGCCATTAGCAAAAGATACAGGATTCGTGACCTACA----- 1260
DB 1831 TCCGGAATATGCCATTAGCAAAAGATACAGGATTCGTGACCTACAAGGTAAGCACAACT 1890
QY 1261 ----- 1260
DB 1891 CTAAGCATACCTTAATGGCTATCTTTCAGAGTATCTGACACAAGAGACTAATCACTGGCA 1950
QY 1261 ----AGAACTGGCCCATCTACAAAGACGACACAACGATGCCATGCGCAAGGGCACAGAT 1316
DB 1951 ATACGAACCTGGCCCATCTACAAAGACGACACAACGATCCCGATGCGCAAGGGCACAGAT 2010
QY 1317 GGTCTCCAGATCGTGAATCTTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCCCTC 1376
DB 2011 GGTCTCCAGATCGTGAATCTTTGTCCAAACAGGGTCTTCGGGTGATTCGTATACCCCTC 2070
QY 1377 TCCTTGAGTGTGCGGGTTACAGCGCGGACCAATGACGAGGTGATTCGTGTCGACG 1436
DB 2071 TCCTTGAGTGTGCGGGTTACAGCGCGGACCAATGACGAGGTGATTCGTGTCGACG 2130
QY 1437 ACCGTACCGTTGGTTCGGATGGAATGTCCCTTCCTATGCGAGTGGCTACCTAGG 1496
DB 2131 ACCGTACCGTTGGTTCGGATGGAATGTCCCTTCCTATGCGAGTGGCTACCTAGG 2190
QY 1497 GTATTGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 1556
DB 2191 GTATTGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG 2250
QY 1557 GAGAGTATATGATGGTACTGCTATTCAATCTGGCATTTGACAGTGTGAGTTGATG 1616
DB 2251 GAGAGTATATGATGGTACTGCTATTCAATCTGGCATTTGACAGTGTGAGTTGATG 2310
QY 1617 TAC 1619
DB 2311 GAC 2313

RESULT 15

US-10-815-495-17
; Sequence 17, Application US/10815495
; GENERAL INFORMATION:
; APPLICANT: Novozymes Biotech, Inc.
; APPLICANT: Connelly, Mariah
; APPLICANT: Brody, Howard
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient
; FILE REFERENCE: 10345.200-US
; CURRENT APPLICATION NUMBER: US/10/815,495
; CURRENT FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 2443

Db	2131	ACCGTGACGGTTGGTTCGGATGGAATGTGCCTGTTCCTATGCGAGGTGGCTACCTAGG	2190
Qy	1497	GTATTGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG	1556
Db	2191	GTATTGTATCCGACTGAGAAGTTGGCAGGTAGCAAGATCTGTAGTAGCTCGTGAAGGGTG	2250
Qy	1557	GAGAGTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGTGAGTTTGATG	1616
Db	2251	GAGAGTATATGATGGTACTGCTATTCAATCTGGCATTGGACAGTGTGAGTTTGATG	2310
Qy	1617	TAC	1619
Db	2311	GAC	2313

Search completed: February 12, 2005, 03:41:54
 Job time : 6093 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2005, 18:00:03 ; Search time 5230 Seconds
(without alignments)
12620.161 Million cell updates/sec

Title: US-10-820-200-1
Perfect score: 1734
Sequence: 1 tcatatcagctctcccttc.....aaaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.8	14.8	583	1	AA786071 j4h07al.r
2	240.2	13.9	541	1	AA787396 n3a12al.r
3	227.2	13.1	561	1	AA783132 c1c11al.r
4	224.2	12.9	854	7	CF686839 CCAE814TR
5	213.4	12.3	937	7	CF711819 CCAEG51TR
6	210.2	12.1	464	1	AA965792 o3e03al.r
7	200.6	11.6	779	7	CF705447 CCAE07al.r
8	189.8	10.9	466	1	AI211322 o6c07al.r
9	184	10.6	753	7	CF709335 CCAA414TR
10	177.8	10.3	783	7	CF679247 CCAAM64TR
11	176.2	10.2	764	7	CF719076 CCAAQ16TR
12	176	10.1	762	7	CF687484 CCAE909TR
13	168.4	9.7	589	7	CN252237 EST018169
14	166.2	9.6	798	7	CF691968 CCAE888TR
15	161.4	9.3	865	7	CF719569 CCAH040TR
16	158.6	9.1	896	7	CF685138 CCAH37TR
17	158.4	9.1	871	7	CF721667 CCAE943TR
18	158	9.1	888	7	CF677145 CCAE42TR
19	156.4	9.0	731	7	CF688600 CCAE001TR
20	155.4	9.0	860	7	CF676000 CCAH72TR
21	155.4	9.0	879	7	CF695401 CCAH059TR
22	152.6	8.8	891	7	CF706510 CCAE320TR
23	151.6	8.7	820	7	CF686770 CCAE632TR
24	151.6	8.7	830	7	CF700769 CCAH804TR

25	151.6	8.7	834	7	CF715868	CF715868	CCAGT30TR
26	151.6	8.7	842	7	CF718519	AA785819	CCAHY86TR
27	150.8	8.7	347	1	AA785831	h8a04al.r	
28	149.4	8.6	625	7	CO134604	EST829275	
29	146.8	8.5	800	7	CF693233	CCAI432TR	
30	143.2	8.3	376	1	AA786814	m5c11al.r	
31	142.2	8.2	796	7	CF697014	CCAAK22TR	
32	140.2	8.1	374	1	AA785147	g5c10al.r	
33	139.4	8.0	382	1	AA785574	g8h0al.r	
34	139.4	8.0	358	1	AI209683	c1c11al.r	
35	137.4	7.9	630	1	AJ638179	AJ638179	
36	135.6	7.8	805	7	CF709195	CCAAD03TR	
37	135.6	7.8	888	7	CF676554	CCADK28TR	
38	134.6	7.8	809	7	CF683596	CCAC892TR	
39	134.2	7.7	867	7	CF695476	CCAB840TR	
40	133.6	7.7	777	7	CF685049	CCAH818TR	
41	133.2	7.7	566	4	BM869651	mgns007xk	
42	133	7.7	389	1	AI212282	x1f04al.r	
43	132.6	7.6	770	7	CF713048	CCAI083TR	
44	131.8	7.6	358	1	AA784386	d5b11al.r	
45	131.2	7.6	304	1	AI210530	i7h08al.r	

ALIGNMENTS

RESULT 1
AA786071
LOCUS
DEFINITION
AA786071 583 bp mRNA linear EST 31-JUL-1998
j4h07al.r1 Aspergillus nidulans 24hr asexual developmental and
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
j4h07al 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA786071 583 bp mRNA linear EST 31-JUL-1998
j4h07al.r1 Aspergillus nidulans 24hr asexual developmental and
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
j4h07al 5', mRNA sequence.
AA786071.1 GI:2846239
EST.
Emericella nidulans (anamorph: Aspergillus nidulans)
Emericella nidulans
Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
1 (bases 1 to 583)
Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
Prade,R. and Roe,B.
An Aspergillus nidulans EST Database
Unpublished (1998)
Other ESTs: j4h07al.f1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: SK
High quality sequence stop: 526.
Location/Qualifiers
1..583
/organism="Emericella nidulans"
/mol_type="mRNA"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="j4h07al"
/tissue type="vegetative mycelia, 24hr asexual
developmental and vegetative cDNA lambda zap library"
/note="Vector: pBluescript SK-; Site1: EcoRI; Site2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN

Query Match 14.8%; Score 256.8; DB 1; Length 583;
Best Local Similarity 65.2%; Pred. No. 1.2e-59;

AUTHORS Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R., and Roe, B.
TITLE An *Aspergillus nidulans* EST Database
JOURNAL Unpublished (1998)
COMMENT Other ESTs: c1clal.f1
 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
 Seq primer: SK
 High quality sequence stop: 511.

FEATURES
 source
 1..561
 /organism="Emericella nidulans"
 /mol_type="mRNA"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="c1clal1"
 /tissue_type="vegetative mycelia, asexual structures"
 /clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
 /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN
 Query Match 13.1%; Score 227.2; DB 1; Length 561;
 Best Local Similarity 68.1%; Pred. No. 1.9e-51;
 Matches 316; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 129 TGGCGATCGCAATCCATTTATTTCTCTCAGCATCGATTTGCAAGGAGGATGGTCG 188
 |||||
 DB 98 TGGCGAGCGCAGTCGATCTATTTCTCTCAGCGAGCGATTCGTCGACGAGCGATCG 157
 |||||
 QY 189 ACCACTGCGACTTGTATATCTGCGGATCAGAAATCTGTGGTGGAACTGGCAGGCGATC 248
 |||||
 DB 158 ACACCGCTCCCTGTGATCTGGCTCAACGAGATATCTGCGCGCGGCGAGCTGGAGGGGATC 217
 |||||
 QY 249 ATCGACAAGTTGGAATATATCCAGGGAAATGGGTTTACAGCCATCTGGATCACCCCGTT 308
 |||||
 DB 218 ATCAACCGCTGATATATCCAGACATGGGATTCATGCCATCTGGATTCACCTATC 277
 |||||
 QY 309 ACAGCCGAGTGCCTCCAGACACCGCATATGGAGATGCCTACCATGGCTACTGGCAGCAG 368
 |||||
 DB 278 ACCGAGCAGATTCGCGATGTACCGCTGTGGAAACGGGCTTCCATGGCTACTGGCAGAG 337
 |||||
 QY 369 GATATATCTCTGAAACGAAACTACGGCACTGCAGATGACTTTGAAGGGCTCTCTTCG 428
 |||||
 DB 338 AACATATACGGTGTGACACCACTGGGACAGCCGACGATATCAGGGCTCTGTGGAG 397
 |||||
 QY 429 GCCCTTTCATGAGAGGGGATGTATCTTTATGGTCGATGTGGTGTCTTAACCATATGGGCTAT 488
 |||||
 DB 398 GCGCTCATGATCGGGCATGTATCTCATGCTGGATGTCTTGGCAACCATGCTTAT 457
 |||||
 QY 489 GATGGAGCGGTAGCTCAGTCGATATCAGTGTGTTTAAACCGTTAGTTCCCAAGATAC 548
 |||||
 DB 458 GGGCGGCCCGCGGATCAACTGACTTCAGATATTTACCCCGTTTCGACTCGGGCTCTAC 517
 |||||
 QY 549 TTCACCCGTTCTGTTTTCATTCAAAACTATGAAGATCAGACTCA 592
 |||||
 DB 518 TTCCATTCGTACTCGGCATCAACACTATGACATCAGTGGCA 561
 |||||

RESULT 4
 CF686839 854 bp mRNA linear EST 16-AUG-2004
 LOCUS C-neoformans strain JEC21 Cryptococcus neoformans var.
 DEFINITION neoformans cDNA clone CCAEE14, mRNA sequence.

CF686839 GI:41540998
 CF686839.1
 EST.
 Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
 Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.
 1 (bases 1 to 854)
 Loftus, B.
 End sequencing of clones from a Full length enriched, normalized
 JEC21 cDNA library
 Unpublished (2003)
 Other ESTs: CCAEE14TO
 Contact: Brendan Loftus
 TIGR
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: crypt@tigr.org
 Seq primer: TR

FEATURES
 source
 1..854
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCAEE14"
 /clone_lib="C-neoformans strain JEC21"
 /note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN
 Query Match 12.9%; Score 224.2; DB 7; Length 854;
 Best Local Similarity 54.9%; Pred. No. 1.4e-50;
 Matches 464; Conservative 0; Mismatches 378; Indels 3; Gaps 1;

QY 247 TCATCGACAAGTTGGACTATATCCAGGGAATGGGCTTCAGGCATCGATCACCCCG 306
 |||||
 DB 5 TCGCCACAACTCGACTATATTCAGGGAATGGGCTTCGACATCTGTGGATATCTCCG 64
 |||||
 QY 307 TTACAGCCGAGTGCCTCCAGACACCGCATATGAGATGCTTACCATGGCTACTGGCAGC 366
 |||||
 DB 65 TCGTGAACAACATGGCGGCACCTACAGGCGAGGCGAGGATATCAGGTTACTGGACTC 124
 |||||
 QY 367 AGGATATATCTCTTGAAACGAAACTACGGCATCGCATGACTTTGAAGGCGCTCTCTT 426
 |||||
 DB 125 TTGATCCCGAGTCCCTCAACTCCAACTTTGGGTCGCTGATGACCTTAAGACTTTGAGTT 184
 |||||
 QY 427 CGGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGGTGTGTCACCATATGGGCT 486
 |||||
 DB 185 CTTCAATTGACGCCAGGGAATGTACTCCAGGTCGACGTCGCTTAACACCATGCTG 244
 |||||
 QY 487 ATGATGGAGCGGTAGCTCAGTCGATTTACAGTGTGTTTAAACCGTTTCAAGTTCACAGACT 546
 |||||
 DB 245 CCAGCTCTAGTTCTACCTTCCAAACCTCCGAATCTTACGGTCTCTTCAGTACTCTGAAG 304
 |||||
 QY 547 ACTTCACCCGTTTCTGTTTTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGGATGCT 606
 |||||
 DB 305 ACTTACCACCTTCTGCTGGATCAGTACTACAAACACAGACCAACGTTGAGCAATGTT 364
 |||||
 QY 607 GGCTAGGAGATACACTGTCTCTGCTGATCTCGATACCAACAGGATGTGGTCAAGA 666
 |||||
 DB 365 GGCTTGGTACGACTCTGTGCTTTTGGCGGACTTGAACACTGAGAGTGACACTGTGTTT 424
 |||||
 QY 667 ATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGGCTCCGTA 726
 |||||
 DB 425 CTTACTGGAAACCAATGATACAGACCTTGTTCCAACTACACTTTTCGATGCCATTCGAA 484
 |||||

QY 727 TCACACAGTAATAACACGTCAGAGGACTTCTGGCCGGGTACAAACAAAGCCGAGCG 786
 |||||
 Db 485 TTGATACCGTCAAGCATGTTGTCATCTTCTGGCCAGATTGTCACACGCTGCGGGT 544
 |||||
 QY 787 TGTACTGTATCGCGAGGTGCTCGACGGTATCCGSCCTA---CACTTGCCTTACCAGA 843
 |||||
 Db 545 TCTTCAACAGGGTGAAGTCTCTCGCGCGATGTGCCATATGCTGCTGGCTACCAAGA 604
 |||||
 QY 844 ACCTCATGACGCGTACTGAATATCCATTTACTATCCACTCTCTCAACGCTTCAAGT 903
 |||||
 Db 605 ATGCGTCTATCAACCTTTCAACTACCCGTTTATATCCCTTGTTCGCGTTTCAACG 664
 |||||
 QY 904 CAACTCCGCGACGATGCGACCTCTACACATGATCAACACCGTCAAAATCCGACTGTC 963
 |||||
 Db 665 GCACTGGACAGAGCTTCAGCGAGCTCATCAACATGGTTTCTTATGTAATAGCAACTTCA 724
 |||||
 QY 964 CAGACTCAACATCTCTGGGACATTTGCTCGAGAACCAACGACCCAGGTTGCTTCTT 1023
 |||||
 Db 725 CCGACCTTACCCTTCTTGTGACTTTCTTGGCAACACACGATAACGCGAGTTTGAGAGCT 784
 |||||
 QY 1024 ACACCAACACATAGCCCTCGCCCAAGACGTCGCGACATTTCATCTCAACGACGGAA 1083
 |||||
 Db 785 TTGTACCCACCTTCATTGATCAAGATGCTCAGGCTTATCTCTTGTCAACGATGGA 844
 |||||
 QY 1084 TCCCC 1088
 |||||
 Db 845 TTCCC 849
 |||||

RESULT 5
 CF711819
 LOCUS
 DEFINITION C.neoformans strain JEC21 Cryptococcus neoformans var.
 neoformans cDNA clone CCAEG51, mRNA sequence.

ACCESSION CF711819
 VERSION CF711819.1 GI:41565978
 SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
 var. neoformans)

ORGANISM Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.

REFERENCE 1 (bases 1 to 937)

AUTHORS Loftus, B.

TITLE End sequencing of clones from a Full length enriched, normalized

JOURNAL JEC21 cDNA library

COMMENT Unpublished (2003)

Contact: Brendan Loftus

TIGR

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: crypt@tigr.org

Seq primer: TR.

FEATURES Location/Qualifiers

1..937

/organism="Cryptococcus neoformans var. neoformans"

/mol_type="mRNA"

/strain="JEC21"

/db_xref="taxon:40410"

/clone="CCAEG51"

/clone_lib="C.neoformans strain JEC21"

/note="Vector: pCMVSPORT6; Site 1: NotI BcoRV; The full

length, normalized library was prepared from a variety of

conditions using RNA provided by Joseph Heitman and

Jennifer Lodge"

ORIGIN

Query Match 12.3%; Score 213.4; DB 7; Length 937;

Best Local Similarity 55.4%; Pred. No. 1.4e-47;

Matches 456; Conservative 0; Mismatches 361; Indels 6; Gaps 2;

QY 121 CTGCGACTGCGCATGCAATCCATTATTTCTTCTCACGGATCGATTTGCAAGACGG 180
 |||||
 Db 102 CTGATCGCATGCGCTCGGCTCAGTCTACCAAGTCGTCAACGATCGATTTGCGCGTA --- 158
 |||||
 QY 181 ATGGGTGACGACTCGGACTTGTAAATATCTGGGATCAAGAAATATCTGTGTGTAACATGCG 240
 |||||
 Db 159 ATTCTTCTTCAACCGCGATTTGTATGTGCGCGACCGCACATATCTGCGGTGCTTGTGT 218
 |||||
 QY 241 AGGCATCATGACAAAGTTGGACTATATCCAGGGAATGGGCTTTCACAGCATCTCGATCA 300
 |||||
 Db 219 CGCTCTCGCGACAAACTCGACTATATTCAGGGAATGGGCTTTCACACTGTCTGGATAT 278
 |||||
 QY 301 CCCCCGTTACAGCCAGCTGCCCCAGACACCGCATATGGAGATCCCTACCATGCTACT 360
 |||||
 Db 279 CTCGGTCTGTCGAAAACATTTGGCGCACTACAGGCGAGGCGAGGCATATCAGGTTACT 338
 |||||
 QY 361 GCGACAGGATATATATCTCTGAAACGAAACTACGGCACTGACAGATGACTTTGAAGGCG 420
 |||||
 Db 339 GGACTTTTATCCCGAGTCCCTCAACTCTTGGGTCGCTGATGACCTTAAGACTT 398
 |||||
 QY 421 TCTCTTCGCGCTTTCATGAGAGGGGATATCTTATGTCGATGTTGTTGTAACCATTA 480
 |||||
 Db 399 TGAGTTCTTTCATGACGCCAGGGAATGTACTCCAGTGCAGCTGCTGCTTAACCG 458
 |||||
 QY 481 TGGGCTATGATGGAGCGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCC 540
 |||||
 Db 459 TCGTGCACGCTCTAGTTCTACTTCCAAACCTCCGAATCTTACGGTCTTTTCAGTACTT 518
 |||||
 QY 541 AAGACTACTTCCACCCGTTCTGTTTCATTCATAAACTATGAAGATCAGACTCAGGTTGAGG 600
 |||||
 Db 519 CTGAAGACTTACCACTCTTCTGCTGGATCACTGACTACAAACCAACAGACCAACGTTGAGC 578
 |||||
 QY 601 ATTGCTGCTAGGATAAACAATGCTCTCTGCTGATCTCGATACCAACAGGATGTCG 660
 |||||
 Db 579 AATGTTGGCTTGGTGACGACTCTGTCGCTTTGGCGGACTTGAACACTGAGATGACACTG 638
 |||||
 QY 661 TCAAGAAATGATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTTCCATTTGACGGCC 720
 |||||
 Db 639 TGTGTTCTTACTGGAACCAATGGATCACAGACTTGTTCACAACTACACTTTTCGATGCCA 698
 |||||
 QY 721 TCCGTATCGACACAGTAAACACAGTCCAGAGGACTTCTGGCCCGGGTACAAACAGCG 780
 |||||
 Db 699 TTGGAATTTGATACCTGCAAGCATGTTCTGCAATCTTCTGCGCAGATTTGCTCAACGCTG 758
 |||||
 QY 781 CAGGCGTCTACTGTATCGCGAGGTGCTCGAGGTGATCCGCGCTA---CACTTGCCT 837
 |||||
 Db 759 CGGTGCTTCAACACAGGGTGAAGTCTCTCGCGGATGTGCCATATGTCGCTGCTACC 818
 |||||
 QY 838 ACCAGACGTCATGGACGCGGTACTGAACTATCCCATTTACTTCCACTCTCAACGCT 897
 |||||
 Db 819 AGAAGATGCGTCTATCAACCTTTCACTACCCCGTTTATATCCCTTGTTCGCGCTT 878
 |||||
 QY 898 TCAAGTCAACCTTCGCGCAGCATGAGACCTTCAACATGAT 940
 |||||
 Db 879 TCAACGGCACTGGACAGAGCTTCAGCGAGCTCATCAACATGCT 921
 |||||

RESULT 6
 AA965792
 LOCUS
 DEFINITION o5e03a1.r1 Aspergillus nidulans 24hr asexual developmental and
 vegetative cDNA lambda zap library Emericella nidulans cDNA clone
 o5e03a1 5', mRNA sequence.

ACCESSION AA965792
 VERSION AA965792.1 GI:3139676
 KEYWORDS EST.
 SOURCE Emericella nidulans (anamorph: Aspergillus nidulans)

ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; Emericella.
 REFERENCE 1 (bases 1 to 464)
 AUTHORS Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,

TITLE		Prade,R. and Roe,B.	
JOURNAL		An Aspergillus nidulans EST Database	
COMMENT		Unpublished (1998)	
		Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu	
		Department of Chemistry and Biochemistry	
		Advanced Center for Genome Technology, University of Oklahoma	
		620 Farrington Oval, Norman, OK 73019, USA	
		Tel: 405 325 4912	
		Fax: 405 325 7762	
		Email: broe@ou.edu	
		We anticipate the future release of the cDNA clones to the Fungal	
		Genetics Stock Center	
		Seq primer: T3	
		High quality sequence stop: 453.	
FEATURES		Location/Qualifiers	
source		1. .464	
		/organism="Emericella nidulans"	
		/mol_type="mRNA"	
		/strain="FGSC A26"	
		/db_xref="taxon:162425"	
		/clone="o5e03a1"	
		/tissue_type="vegetative mycelia, asexual structures"	
		/clone_lib="aspergillus nidulans 24hr asexual	
		developmental and vegetative cDNA lambda zap library"	
		/notes="vector: pBluescript SK-; Site 1: EcoRI; Site 2:	
		XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript	
		3' end of cDNA cloned into XhoI site of pBluescript"	
ORIGIN			
Query Match		12.1%; Score 210.2; DB 1; Length 464;	
Best Local Similarity		65.9%; Pred.No. 9.4e-47;	
Matches 305; Conservative		0; Mismatches 158; Indels 0; Gaps 0;	
QY	820	CGGCCTACACTGTGCCCTACGAGACGTGATGACGGCGTACTGAACCTATCCCATTTACT	879
DB	1	CGGCGTATACCTGCGCCTTACAGATTATATGACGGGGTCATGAACCTACCCCATATATT	60
QY	880	ATCCACTCTCTCAACGGCCTTCAAGTCAACCTCCGGCAGCATGGACGACTCTACAACATGA	939
DB	61	ACCCCTTCTGAATGCGTTCAAGTCTCGAGCGGAGCATGTCGGATCTCTATAACATGA	120
QY	940	TCAACACCGTAAATCGCACTGTCCAGACTCAACACTCTCGGGGCACATTCGTGAGAAC	999
DB	121	TCAACACAGTCGCGCTCAAATTTGTGGGATCCTACACTGCTTGGAACTTTATCGAGAAC	180
QY	1000	ACGACACCCACGGTTCGCTTCTTACACCAACGACATAGCCCTCGCCAAAGAGCTGCGAG	1059
DB	181	ATGACAACTCTCGAATCCCAACTATCTACTCCGATATAGTCGGGCCAAGAGCTCTCTCG	240
QY	1060	CATTTCATCTCTCAACGACGGGAATCCCATCATCTACGCGCGGCCAAGAACGACCTACG	1119
DB	241	CGTTCTCTTCTTGACCGCAGGAATCCCTATTTGTTATGCGGGCCAGGACGACACTATT	300
QY	1120	CGGCGGAACGACCCCGGAAACCGGAAAGCAACCTGGCTCTCGGGGTACCCGACCGACA	1179
DB	301	CAGGCGCAATGATCCCTATTAACCGGAGCGGTTTGGTGGTCTCTACTCGACCAAGCT	360
QY	1180	GCAGCTGTACAGTTAATTGGCTCCGCGAACCAATCCGGAACCTATGCCATTAGCAAG	1239
DB	361	CAGAGCTATACAAGTTTCATCGGACCACTAACAGATCCGAAACTGGCCATTTCCAAG	420
QY	1240	ATACAGGATTCTGTGACCTACAAGAACTGGGCCCATCTACAAGA	1282
DB	421	ATTTCCAGTTATCTCACTTCCGGNACTCTCTTTTACACGA	463
RESULT 7			
CF705447			
LOCUS		779 bp mRNA linear EST 16-AUG-2004	
DEFINITION		C.neoformans C.neoformans strain J8C21 Cryptococcus neoformans var.	
ACCESSION		neoformans cDNA clone CCABA08, mRNA sequence.	
VERSION		CF705447	
		CF705447.1 GI:41559606	

KEYWORDS	EST.
SOURCE	Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
ORGANISM	Cryptococcus neoformans var. neoformans Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
REFERENCE	1 (bases 1 to 779)
AUTHORS	Loftus,B.
TITLE	End sequencing of clones from a Full length enriched, normalized JEC21 cDNA library
JOURNAL	Unpublished (2003)
COMMENT	Other_ESTs: CCABA08TF Contact: Brendan Loftus TIGR
	Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-3543 Fax: 301-838-0208 Email: crypt@tigr.org Seq primer: TR.
FEATURES	Location/Qualifiers
source	1..779 /organism="Cryptococcus neoformans var. neoformans" /mol_type="mRNA" /strain="JEC21" /db_xref="taxon:40410" /clone="CCABA08" /clone_lib="C.neoformans strain JEC21" /note="Vector: pCMVSPORT6; Site_1: NotI_EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"
ORIGIN	Query Match 11.6%; Score 200.6; DB 7; Length 779; Best Local Similarity 55.5%; Pred. No. 4.8e-44; Matches 430; Conservative 0; Mismatches 339; Indels 6; Gaps 2; QY 157 TCACGGATCGATTGCAAGGCGGATGGGTGCGAGCTCGGACTTGTAACTATCTGCGGATC 216 DB 8 TCACCGATTGATTGGCGGTA---ATTCTTCTTCAACGGCGGATTGAATGTGCGCGACC 64 QY 217 AGAAATACTGTGTGGTGAACATGGCAGGGGCATCATCGACAAGTTGGACTATATCCAGGAA 276 DB 65 GCACATACTCGGTTGGTACTTGGTCGGCTCTCGCGACAACTCGACTATATTCAGGGAA 124 QY 277 TGGGCTTCAAGCCATCTGGATCAACCCCGTTACAGCCAGCTGCCCCAGACACCAGCAT 336 DB 125 TGGGCTTTGACACTGTCTGGATATCTCCGGTCGTGAAACATTTGGCGGCACTACAGGCG 184 QY 337 ATGGAGATCGCTACCATGCTACTGGCAGCAGGATATATCTCTGTAACGAAACTAGC 396 DB 185 AGGGGAGGCAATATCACGGTTACTGGACTCTTGAATCCCGAGTCCCTCAACTCCACTTTG 244 QY 397 GCATCTCAGATGACTTGAAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTAA 456 DB 245 GGTCCGCTGATGACCTTAAGACTTTGAGTTCTTCATTGACGCCAGGGAAGTACCTCC 304 QY 457 TGGTCGATGTGGTTGCTAAACCATATGGGCTATGATGGAGCGGGTAGCTCAGTCGATTACA 516 DB 305 AGGTCGACGCTCGTTAACCAACGCTGCTGCCACGCTAGTTCTTACCTTCCAACCCCTCG 364 QY 517 GTGTGTTTAAACGGTTCAGTCTCCCAAGACTACTTCCACCGGTTCTGTTTCATTCAAAACT 576 DB 365 AATCTTACGGTCTCTTTCAGTACTTCTGAAAGACTACACCCCTTTCTGCTGGATCACTGACT 424 QY 577 ATGAAGATCAGACTCAGGTTGAGGATGTGCTGGCTAGGAGATAAACACTGTCTCTTGCCTG 636 DB 425 ACACACACGACCAACGTTGAGCAATGTTGGCTTGGTGACACTCTGTCTGCTTTGGCCG 484 QY 637 ATCTCGATACCAACGAGGATGTGGTCAAGAAATGAATGGTACGACTGGGTGGGATCATTTGG 696

```

Db      485 ACTGAACACTGAGAGTGACATGTTGTTTCTTACTGGAAACCAATGGATCAACAGACCTTG 544
QY      697 TATCGAACTACTCCATTTGACGGCTCCGATCGACACAGTAAACACGTCACAGAGGACT 756
Db      545 TTTCCAACTACACTTTCGATGCCATTCGAATGATACCGTCAAGCATGTTGCTCAATCTT 604
QY      757 TCTGGCCGGGTACAAACGAGCGGAGGCGGTGTACTGTATCGCGGAGGTGCTCGACGGTG 816
Db      605 TCTGGCCAGATTTCGTCAACGGCTGCGGTGTCTTCAACACGAGGTGAAGTCTCTCTGGCGG 664
QY      817 ATCCGGCCTA---CACTGTCTCCTACCAGAAAGTCAATGACCGGCTACTGAATATCCCA 873
Db      665 ATGTGCATATGTGCTGTGCTACCAAGAAATGCGTCTATCAACCTTTCAACTACCCCG 724
QY      874 TTTACTATCCACTCTCTCAACGGCTTCAAGTCAACCTCGGACGATGGACGACCT 928
Db      725 TTTATTAATCCCTTGTTCGCGTTTCAACGGCACTGGACAGAGCTTCACGGAGCT 779

RESULT 8
LOCUS   AI211322                466 bp mRNA linear EST 19-OCT-1998
DEFINITION o6c07al.r1 Aspergillus nidulans 24hr asexual developmental and
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
o6c07al 5', mRNA sequence.
ACCESSION AI211322.1 GI:3773264
VERSION    AI211322
KEYWORDS   Aspergillus nidulans
SOURCE     Emericella nidulans (anamorph: Aspergillus nidulans)
ORGANISM   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
REFERENCE  1 (bases 1 to 466)
AUTHORS    Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
Prade,R. and Roe,B.
TITLE      An Aspergillus nidulans EST Database
JOURNAL    Unpublished (1998)
COMMENT    Other ESTs: o6c07al.f1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: T3
High quality sequence stop: 321.

FEATURES             source
    Location/Qualifiers
        1..466
            /organism="Emericella nidulans"
            /mol_type="mRNA"
            /strain="FGSC A26"
            /db_xref="taxon:162425"
            /clone="o6c07al"
            /tissue_type="vegetative mycelia, asexual structures"
            /clone_lib="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
            /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN
Query Match      10.9%; Score 189.8; DB 1; Length 466;
Best Local Similarity 70.4%; Pred. No. 4.3e-41;
Matches 254; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY      129 TGGCGATCGCAATTCATTTATTTCTTCACGGATCGATTTGCAAGGACGATGGGTGG 188
Db      104 TGGCGAGCGAGTCGATCTATTTCTTCCTCGACGACCGATTCGCTCGACGACGATCG 163
QY      189 ACGACTGCGACTTGTAAATCTCGGATCAGAAATACTGTGTGGAAACATGGCGGGCATC 248

```

```

Db      164 ACGACCGCTGCGTGTGATCTGGCTCAACGGAGATCTGCGGCGGAGCTGGCAGGGCATC 223
QY      249 ATCGACAAGTTGGACTATATCCAGGGAATGGCTTTCACAGCCATCTGGATCACCCCGTT 308
Db      224 ATCAACCACTGTGATTATATCAAGACATGGATTCTAGCCATCTGGATTACACCTATC 283
QY      309 ACAGCCCACTGCTCCCCAGACACCGCATATGGAGATGCTTACCATGGTCTACTGGCAGCAG 368
Db      284 ACCGAGCAGATTCCCGATGTCAACGCTGTGGAAACGGCTTACATGGCTACTGGCAGAG 343
QY      369 GATATATATCTCTGAACGAAAACTACCGCACTGCAGATGACTTGAAGCGCTCTCTTCG 428
Db      344 AACATATACGGTGTGACACCAACCTGGGSCACAGCCGATATCAGGCTCTGTGCGGAG 403
QY      429 GCCCTTATGAGAGGGGATGTATCTTATGTCGATGTTGCTAACCATATGGCTAT 488
Db      404 GCGCTCCATGATCGGGGCGATGTATCTCATGCTGGATGTTGTGCGAACCATGCTTAT 463

QY      489 G 489
Db      464 G 464

RESULT 9
LOCUS   CF709355                753 bp mRNA linear EST 16-AUG-2004
DEFINITION CCAA414TR C neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCAA414, mRNA sequence.
ACCESSION CF709355
VERSION    CF709355.1 GI:41563514
KEYWORDS   Cryptococcus neoformans var. neoformans
SOURCE     Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
ORGANISM   Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
REFERENCE  1 (bases 1 to 753)
AUTHORS    Loftus,B.
TITLE      End sequencing of clones from a Full length enriched, normalized
JEC21 cDNA library
JOURNAL    Unpublished (2003)
COMMENT    Other ESTs: CCAA414TF
Contact: Brendan Loftus
TIGR
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: cryptot@tigr.org
Seq primer: TR
Location/Qualifiers
    1..753
        /organism="Cryptococcus neoformans var. neoformans"
        /mol_type="mRNA"
        /strain="JEC21"
        /db_xref="taxon:40410"
        /clone="CCAA414"
        /clone_lib="C.neoformans strain JEC21"
        /note="Vector: pCMVSPORT6; Site 1: NotI; EcoRV; The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jennifer Lodge"

ORIGIN
Query Match      10.6%; Score 184; DB 7; Length 753;
Best Local Similarity 56.1%; Pred. No. 1.9e-39;
Matches 368; Conservative 0; Mismatches 285; Indels 3; Gaps 1;

QY      121 CTGCGGATGGGATCGCAATCCATTTATTTCTTCTACGGATCGATTTGCAAGCGG 180
Db      96 CTGATGCGATGGCTCGCGCTCAGTCTACCAAGTCGTCAACCGATCGATTTGCGCGTA--- 152

```



```
QY 181 ATGGGTGCGAGCTCGACTTGTAACTACGCGATCAGAAATCTGTGGTGGNACATGCG 240
Db 153 ATTCTTCTTCAACCGCGGATGTAATGTGCGCGACCGCACATACCTCGGTGGTACTTGT 212
QY 241 AGGGCATCATCGCAAGTGGACTATATCCAGGGAATGGGCTTCCAGAGCATCTGGATCA 300
Db 213 CGCTCTCGCGCAAACTCGACTATATTTCAGGGAATGGCTTTCAGACTGCTGGATAT 272
QY 301 CCCCCGTTACAGCCAGCTGCCCCAGACCAACCGCATATGAGATGCTTACCATGGTACT 360
Db 273 CTCGGGTCTGCAAAACATTTGGCGGCACTACAGCGAGGCGAGGCATATCACGGTTACT 332
QY 361 GCGAGCAGATATATCTCTCTGAACGAAACTACGGCACTCGAGATGACTTGAAGCGC 420
Db 333 GGACTTTGATCCGAGTCCCTCAACTCCCAACTTTGGGTCGCGTGAACCTTAAGACTT 392
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGTCGATGTGGTTCGTAACCAT 480
Db 393 TGAGTTCTTCAATGACGCGGAGGATGTACTTCAGGTGCGACGTGCTGTTAACCAAG 452
QY 481 TGGGTATGATGAGCGGGTAGCTCAGTCGATTTACAGTGTGTTTAAACCGTTCAAGTCCC 540
Db 453 TCGTCCGACGCTAGTCTTACCTTCCAAACCTCCGAATCTTACGGTCTCTTTCAGTACT 512
QY 541 AAGACTACTTCCACCGTTCTGTTTCATTTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
Db 513 CTGAAGACTACCAACCTTTCTGCTGATCACTGACTACAAACCAACAGACCAACGTTGAGC 572
QY 601 ATTGCTGGCTAGGAGATAACACTGTCTCTGCTGATCTCGATATCCCAAGGATGTGG 660
Db 573 AATGTTGGCTTGTGACGACTCTGTCGCTTTGGCGCACTTGAACACTGAGAGTGACACTG 632
QY 661 TCAAGAATCAATGTTACGACTGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGCC 720
Db 633 TTGTTTCTTACTGGAACCAATGATCAGACAGACCTTGTTCCTTCCAACTACACTTCGATGCCA 692
QY 721 TCGTATCGACAGTAAACAGCTCCAGAGGACTTCTGGCCGGGTACAACAAA 776
Db 693 TTGCAATTGATACCGTCAAGCATGTTTCGTCATCTTTCTGGCCAGATTTTCGTCAAA 748
```

```
RESULT 10
CF679247 783 bp mRNA linear EST 16-AUG-2004
LOCUS CCAAM64TR C.neoformans strain JEC21 Cryptococcus neoformans var.
DEFINITION neoformans cDNA clone CCAAM64, mRNA sequence.
ACCESSION CF679247
VERSION CF679247.1 GI:41533406
SOURCE EST.
Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
1 (bases 1 to 783)
Loftus,B.
End sequencing of clones from a Full length enriched, normalized
JEC21 cDNA library
Unpublished (2003)
Other ESTs: CCAAM64TF
Contact: Brendan Loftus
TIGR
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: crypt@tigr.org
Seq primer: TR.
Location/Qualifiers
1..783
/organism="Cryptococcus neoformans var. neoformans"
```

```
/mol_type="mRNA"
/strain="JEC21"
/db_xref="taxon:40410"
/clone_lib="CCAAM64"
/Note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jennifer Lodge"
```

ORIGIN

```
Query Match
Best Local Similarity 10.3%; Score 177.8; DB 7; Length 783;
Matches 360; Conservative 0; Mismatches 282; Indels 3; Gaps 1;
QY 121 CTGCGGACTGGCGATCGCAATCCATTTATTTCTTCTCAGGATCGATTGTCGAAGACGG 180
Db 108 CTGATGCGATGCGCTCGGCTCAGTCTACCAAGTGTCTACCGATGATTTCGCGGTA--- 164
QY 181 ATGGGTGCGAGCTGGACTTGTATACTGCGGATCAGAAATACTGTGGTGGNACATGCG 240
Db 165 ATTCTTCTTCAACCGCGGATTTGTAATGTGCGCGACCGCATATCTGCGGTGGTACTTGT 224
QY 241 AGGGCATCATCGCAAGTGGACTATATCCAGGGAATGGGCTTTCAGAGCATCTGGATCA 300
Db 225 CCGCTCTCGCGCAAACTCGACTATATTTCAGGGAATGGGCTTTCAGACTGCTGGATAT 284
QY 301 CCCCCGTTACAGCCAGCTGCCCCAGACCAACCGCATATGAGATGCTTACCATGGTACT 360
Db 285 CTCGGTCTGCGAAACATTTGGCGGCACTACAGGCGAGGCGGCGCATATCACGGTTACT 344
QY 361 GCGAGCAGATATATCTCTGAAAGAAACTACGGCACTCGAGATGACTTGAAGCGC 420
Db 345 GGACTCTTGTATCCCGAGTCCCTCAACTCCAACTTTGGGTCGCTGATGACCTTAAGACTT 404
QY 421 TCTCTTCGGCCCTTCATGAGAGGGGATGTATCTTATGTCGATGTGGTTCGTAACCAT 480
Db 405 TGAGTTCTTCAATGACGCGGAGGATGTACTTCAGGTGCGACGTGCTGTTAACCAAG 464
QY 481 TGGGTATGATGAGCGGGTAGCTCAGTCGATTTACAGTGTGTTTAAACCGTTCAAGTCCC 540
Db 465 TCGCTGCCAGCTAGTCTTCTACCTTCCAAACCTCCGAATCTTACGGTCTCTTTCAGTACT 524
QY 541 AAGACTACTTTCACCGGTTCTGTTTCATTTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
Db 525 CTGAAGACTATCACACCTTTCTGCTGGATCACTGACTACAAACCAACAGACCAACGTTGAGC 584
QY 601 ATTGCTGGCTAGGAGATAACACTGTCTCTGCTGATCTCGATACCAACAGGATGTGG 660
Db 585 AATGTTGGCTTGTGACGACTCTGCGCTTTGGCGCACTTGAACACTGAGAGTGACACTG 644
QY 661 TCAAGAATCAATGTTACGACTGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGCC 720
Db 645 TTGTTTCTTACTGGGCAACATGATCAGAGACCTTGTTCCTCAACTACACTTTCGATGCCA 704
QY 721 TCGGTATCGACAGTAAACAGCTCCAGAGGACTTCTGGCCCG 765
Db 705 TTGCAATTGATACCGTCAAGCATGTTTCGTCATCTTTCTGCGCAG 749
```

RESULT 11

```
CF719076 764 bp mRNA linear EST 16-AUG-2004
LOCUS CCAAQ16TR C.neoformans strain JEC21 Cryptococcus neoformans var.
DEFINITION neoformans cDNA clone CCAAQ16, mRNA sequence.
ACCESSION CF719076
VERSION CF719076.1 GI:41573235
KEYWORDS EST.
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
```

QY	721	TCGGTATCGACACAGTAAACACGTCAGAGAGACTTCTGCGCG	765
Db	701	TTTCGATTGATACCGTCAAGCATGTTCTCAATCTTTCTGCGCAG	745
RESULT 12			
CF687484			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 359; Conservative			
QY	121	CTGCGACTGGCGATCGCAATCCATTTATTTCTTCTCAGCGATCGATTTGCAAGGACGG	180
Db	104	CTGATCGATGGCTCGCGCTCAGTCTACCAAGTCTGCCGATCGATTTGCGCGTA---	160
QY	181	ATGGGTTCGACGACTGCGACTTGTAACTACTGCGGATCAGAAATACTGTGGTGGAACTGGC	240
Db	161	ATTCTTCTTCAACCGCGGATGTATGTGCGCGACCGCACATACTGCGGTGGTACTTGGT	220
QY	241	AGGGCATATCGACAGTTGGACTATATCCAGGGAATGGGCTTCACAGCCATCTGGATCA	300
Db	221	CGGCTCTCGCCGACAACTCGACTATATTCAGGGAATGGGCTTTGACACTGTCTGGATAT	280
QY	301	CCCCGTTACACCCGACTGCCCCAGACCGCATATGGAGATGCCCTACCATGGCTACT	360
Db	281	CTCCGGTCTCGAAACATTTGGCGGCATCAGCGGAGGCGGAGGCATATCACGGTTACT	340
QY	361	GGCAGCAGGATATATCTCTCTGAAACGAAACTACGGCACTCGACATGACTTGAAGCGC	420
Db	341	GGACTCTTGATCCCGAGTCCCTCACTTCCAACTTTGGTCCGCTGATGACTTAAGACTT	400
QY	421	TCTCTTGGCCCTTCATGAGAGGGGATGTATCTTATGGTTCGATGTGGTGGTCTTAACATA	480
Db	401	TGAGTTTCTTCATTCGACGCGAGGGAATGTACCTCCAGGTCGAGCTCGTTCGTTAACACG	460
QY	481	TGGGCTATGATGGAGGGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTCACTTCCC	540
Db	461	TCGCTGCCACGTCTATGTTCTACCTTCCAAACCCCTCCGAATCTTACGGTCCCTTCAGTACT	520
QY	541	AAGACTACTTCCACCCGTTCTGTCTTTCATTAACAACTATGAAGATCAGACTCAGGTTGAGG	600
Db	521	CTGAAGACTACACCCCTTCTGCTGGATCACTGACTACMACACAGACCAACGTTGAGC	580
QY	601	ATTGCTGGCTAGGAGATAAAGTGTCTCTTTCCTTCCTGATCTCGATACCAACCAAGGATGG	660
Db	581	AATGTTGGCTTGGTACGACTCTGTCTGCTTTGGCGGACTTGAACACTGAGAGTGACACTG	640
QY	661	TCAGAAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCATTGACGGCC	720
Db	641	TTGTTTCTTACTGGAAACCAATGGATCAGACACCTTGTTTTCCAACTACACTTTCGATGCCA	700

```

Db      394 TGAGTTCTTTCATTCACGCCAGGGGAATGTACCTCCAGTTCGAGTCGTGCTTAACCAAG 453
QY      481 TGGGCTATGATGAGCGGGTAGCTCAGTCGATACAGTGTGTTTAAACCGCTTCAGTTCC 540
Db      454 TCGCTGCGACGCTAGTTCTACCTTCCAAACCTCCGAATCTTACGGTCCCTTTCAGTACTT 513
QY      541 AAGACTAC-TTCACACCGTTCTGTTTCATTCATAAAGCTATCAAGATCAGACTCAGGTTGAG 599
Db      514 CTGAAGCATACCAACCTTCTGCTGGATCACTGACTACAAACCCAGACCAACGTTGAG 573
QY      600 GATTGCTGGTAGGAGATAACACTGTCTCCTTGCTGATCTCGATACCAACCAAGGATGTG 659
Db      574 CAATGTTGGCTTGGTGACGACTCTGTACATTTGCCCGACTTGAACACTGAGAGTGACACT 633
QY      660 GTCAGAATGAATGTACGACTGGGTGGATCATTTGGTATCGAAGTACTCCATTCAGCGGC 719
Db      634 GTTGTGTTCTTACTTGAACCAATGGATCACAGACCTTGTGTTTCCAACTACACTTTTCGATGCC 693
QY      720 CTCGGTATCGACAGTAAACACGTCACAGAGACTTTCGGCCCGGTACAAACAGCC 779
Db      694 ATTGGAATTGATACCGTCAAGCATGTTGTCATCTTTCTGGACAGATTTCTGTCACAGCT 753
QY      780 GCAGGCGT 787
Db      754 GCCGGTGT 761

RESULT 13
CN252237
LOCUS      589 bp mRNA linear EST 09-APR-2004
DEFINITION EST018169 Mycelium and yeast cells from Paracoccidioides
brasilienis Paracoccidioides brasiliensis cDNA, mRNA sequence.
ACCESSION CN252237
VERSION   CN252237.1 GI:46355983
SOURCE    EST.
ORGANISM  Paracoccidioides brasiliensis
           Paracoccidioides brasiliensis
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
           Onygenales; Mitosporic Onygenales; Paracoccidioides.
AUTHORS   Felipe,M.S.S., Carvalho,M.J.A., Andrade,R.V., Araes,F.B.M.,
           Simoes,I.C., Andrade,E.V., Maranhao,A.O., Torres,F.A.G.:
           Jesuino,R.S.A., Kyaw,C.M., Moraes,L.M.P., Nicola,A., Pereira,M.,
           Silva-Pereira,I., Anjos,D.A.S., Sandes,E.F.O., Inoue,M.K.,
           Walter,M.E.M.T., Soares,C.M.A. and Brigidio,M.M.
TITLE     Metabolic features of Paracoccidioides brasiliensis cell
           differentiation as accessed by transcriptome analysis
JOURNAL   Unpublished (2004)
COMMENT   Contact: Felipe MSS
           Laboratory of Molecular Biology
           Institute of Biology - University of Brasilia
           Campus Universitario, Asa Norte, Brasilia, DF 70910-900, BRA
           Tel: 55 61 307 2423
           Fax: 55 61 349 8411
           Email: msueli@unb.br
           Seq primer: T7 Sequencing primer.
FEATURES
source    1..589
           Location/Qualifiers
           /organism="Paracoccidioides brasiliensis"
           /mol_type="mRNA"
           /strain="Pb01"
           /db_xref="taxon:121759"
           /clone_lib="Mycelium and yeast cells from Paracoccidioides
           brasiliensis"
           /notes="Pb Lambda Zap Express Library"

ORIGIN
Query Match 9.7%; Score 168.4; DB 7; Length 589;
Best Local Similarity 59.0%; Pred. No. 3.9e-35;
Matches 314; Conservative 0; Mismatches 206; Indels 12; Gaps 1;

QY      94 TCGCGCACCTGTTGGCTGCAACCGCTCGGACTGGGATCGCAATCCATTATTTC 153

```

```

Db      57 TCGCCGGACTAGCATTTAGCAGCGGACACCGTCGATTGGAAATCCAGTCGATATATCAGG 116
QY      154 TTCTCAGGATCGATTTGCAAGACGGATGGGTGCGACGACTGCGACTTGTAAATCTGTCGG 213
Db      117 TCTTACGCGACAGATTTGCGAGGACTGATGGTTTCGACTTCGGCAAAATGCGATACGAACG 176
QY      214 ATCAGAAATACTCTGTGTGAACATGCGAGGGGATCATCGACAGTTGGACTATATCCAGG 273
Db      177 AGAACCTCTACTTCGGGAGGATCATGGAGGGGCATAATCAGCAAGTTGGACTATATTAAAG 236
QY      274 GAATGGCTTTCACAGCCATCTGGATCACCCCGCTTACAGCCCAAGCTGCCCCAGACCAACCG 333
Db      237 AATGGGCTTCGATGCGCATCATGATCTCACCAGTCTCGAAAAATGCGAAGTTCGCGTTA 296
QY      334 CATATGGAGATGCTTACCATGGCTACTGGCAGCAGGATATATACTCTTCGAACGAAACT 393
Db      297 GTTACGGGAAGCATACCATGGCTACTGGGTTCAAGATCTATACCAATTTGAAGCAACATT 356
QY      394 AGGCACTGCAGATGACTTGNAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTATC 453
Db      357 TTGGAATCGCGATGATCTTATCGACCTTGGAAAGGAGGTGCATAAGCGGGGATGTACT 416
QY      454 TTATGGTCGATGTGTTGCTTAACCATATGGGCTTATGATGAGCGGGTAG----- 502
Db      417 TGNATGTCGATATCGTCATCAATAACATGCGCTTCATAAGCATGTGATAGAAACCTGCCA 476
QY      503 -CTCAGTCGATTACAGTGTGTTTAAACCGTTCAGTTTCCCAAGACTACTTCCACCCGTTCT 561
Db      477 TTGGCGTGATTACTCTGTCTCAACCCATTAAACCGCAAGGTATTACACCCGTTACT 536
QY      562 GTTTCATTCAAACATGAAGATCAGACTCAGGTGAGGATTCCTGGCTAGG 613
Db      537 TTAAGATTAAAGCACTGGAATTAATTACGATGAGGCGCAAAATTTGTTGGATAGG 588

RESULT 14
CF691968
LOCUS      798 bp mRNA linear EST 16-AUG-2004
DEFINITION CCADW88TR C.neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCADW88, mRNA sequence.
ACCESSION CF691968
VERSION   CF691968.1 GI:41546127
KEYWORDS  EST.
SOURCE    Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
           var. neoformans)
ORGANISM  Cryptococcus neoformans var. neoformans
           Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
           Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
           Filobasidiella.
REFERENCE Loftus,B.
AUTHORS   End sequencing of clones from a Full length enriched, normalized
TITLE     JEC21 cDNA library
JOURNAL   Unpublished (2003)
COMMENT   Other_ESTs: CCADW88TF
           Contact: Brendan Loftus
           TIGR
           Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
           Rockville, MD 20850, USA
           Tel: 301-838-3543
           Fax: 301-838-0208
           Email: crypt@tigr.org
           Seq primer: TR.
FEATURES
source    1..798
           Location/Qualifiers
           /organism="Cryptococcus neoformans var. neoformans"
           /mol_type="mRNA"
           /strain="JEC21"
           /db_xref="taxon:40410"
           /clone="CCADW88"
           /clone_lib="C.neoformans strain JEC21"
           /note="Vector: pCMVSPORT6; Site 1: NotI_EcoRV; The full
           length, normalized library was prepared from a variety of

```

conditions using RNA provided by Joseph Heitman and Jennifer Lodge		End sequencing of clones from a Full length enriched, normalized JEC21 cDNA library	
ORIGIN		JOURNAL	COMMENT
Query Match	9.6%; Score 166.2; DB 7; Length 798;	Unpublished (2003)	
Best Local Similarity	53.9%; Pred. No. 1.7e-34;	Other ESTs: CCAHO40TF	
Matches 364; Conservative	0; Mismatches 308; Indels 3; Gaps 1;	Contact: Brendan Loftus	
QY	592 AGGTTGAGGATTGCTGGCTAGGAGATAACACTGTCTCTCTGCTGATCTCGATATACCACCA 651	TIGR	
Db	4 AGTTGAGCAATGTTGGCTTGTGTGACGACTCTGTGCTTTGGCGGACTTGAACACTGAGA 63	Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA	
QY	652 AGGATGTGTCAGGAATGAATGGTACGACTGGGTGGGATCATTTGGTATCGAACTACTCCA 711	Tel: 301-838-3543	
Db	64 GTGACACTCTGTTTCTTACTGGAACCAATGGATCACAGACCTGTTTTCCAACTACACTT 123	Fax: 301-838-0208	
QY	712 TTGACGGCTTCGTATCGACACAGTAACAGTCAGTCCAGAAAGGACTTCTGGCCGGGTACA 771	Email: crypt@igr.org	
Db	124 TCGATGCCAATTCGAATTTGATACCGTCAAGCATGTTTCGTCAATCTTTCTGGCCAGATTTCG 183	Seq primer: TR.	
QY	772 ACAAGCCGACGCGTGTACTGTATCGGCGAGTGTCTCGACGTGATCGGCTA---CA 828	Location/Qualifiers	
Db	184 TCAACGCTCCGCTGCTTCAACAGGTTGAAGTCTCTCTGGGATGTGCCATATGTCTG 243	1..865	
QY	829 CTTGTCCCTACAGAACGTCATGGACGGGTACTGAACCTATCCCACTTTACTATCACTCC 888	/organism="Cryptococcus neoformans var. neoformans"	
Db	244 CTGGCTACAGAAATGCGTCTATCAACCTTTCAACTACCCCGTTTATTATCCCTTG 303	/mol_type="mRNA"	
QY	889 TCAACGCTTCAAGTCAACCTCGGCGAGTGAAGCACTCTTACACATGATCAACACCG 948	/strain="JEC21"	
Db	304 TTGCGGCTTCAACGGCACTGGACAGAGCTTCAGCGAGCTCATCAACTGGTTTCTTATG 363	/db_xref="taxon:40410"	
QY	949 TCAATTCGACTGTCCAGACTCAACTCTCTGGGCATTTCTGCGAAGACCAACCAACC 1008	/clone="CCAHO40"	
Db	364 TAAATAGCAACTTCAACGACCTTACCCCTTCTTGATTTCTTGGACAACAGATAACG 423	/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"	
QY	1009 CACGGTTCCGCTTTTACACCAACGATAGCCCTCGCCAAAGACGTCGAGCATTCATCA 1068		
Db	424 CCAGGTTTGAGAGCTTTGTTCACCGACCTTCATTGATCAAGAAATGCTCAGGCTTATCCTC 483	ORIGIN	
QY	1069 TCTCAACAGCGAATCCCATCATCTACGCGGCCAAGAACAGCACTACGCGGGCGAA 1128	Query Match	9.3%; Score 161.4; DB 7; Length 865;
Db	484 TTGTACCGATGGAATTCCTATGTTTACTACAGTAGTGAAGCGGCTTCAACGGTGGCA 543	Best Local Similarity	55.9%; Pred. No. 3.7e-33;
QY	1129 AGACCCCGGAACCGCGAAGCAACTGCTCTCGGGCTACCGACGACGAGCTGT 1188	Matches 328; Conservative	0; Mismatches 256; Indels 3; Gaps 1;
Db	544 AGACCTTGAGAACCGAGAGCCCATGTGGACCAAGCACTACGACCACTTCGACATGT 603	QY	121 CTGCGACTGGCGATCGCAATCCATTTATTTCTCTACGGATCGATTTGCAAGACGG 180
QY	1189 ACAAGTTAATGCTCCGGAAGCGAATCCGGAACCTATGCCATTAGCAAGATACAGGAT 1248	Db	282 CTGATGCGATGCGTCGCGCTCAGTCTACCAAGTCGTACCGATCGATTTGCGCGTA--- 338
Db	604 ACAAGTTCTCACTTCTTAAAGCTGCGCGTCTCTGCTGCTGCAACGCTTCTAGCACTT 663	QY	181 ATGGGTGACGACTCGGACTTGTATATCTACGCGATCAGAAATACTGTGTGGAACATGSC 240
QY	1249 TCGTGACCTACAGA 1263	Db	339 ATTCTTCTTCAACCGGCGATTTGTAATGTGCGCGACCGCACATCTACTGCGGTGTACTTGT 398
Db	664 TCTACTACACAAGA 678	QY	241 AGGGCATCATCAAGTTGGACTATATCCAGGGAATGGCTTTCACAGGCATCTGGATCA 300
RESULT 15		Db	399 CCGCTCTGCCGACAACTCGACTATATTCAGGGAATGGGCTTTGACACTGTCTGATAT 458
CF719569	865 bp mRNA linear EST 16-AUG-2004	QY	301 CCCCCTTTACAGCCAGCTGCCCGACACCGCATATGAGATGCCTTACCATGGTACT 360
CCAHO40TR	C.neoformans strain JEC21	Db	459 CTCGGTCTGCGTGAATACTTGGCGGCACTACAGGCGAGGCGGCGCATATCACGGTTACT 518
DEFINITION	neoformans cDNA clone CCAHO40, mRNA sequence.	QY	361 GGCACAGATATATATCTCTGAACGAAAACACGCGACTCGAGATGAGTGTGAAGCGC 420
ACCESSION	CF719569	Db	519 GGACTCTTGATCCCGAGTCCCTCAACTCCAACTTTGGGTCCGCTGATGACCTTAAGACTT 578
VERSION	CF719569.1	QY	421 TCTCTCGGCCCTTCATGAGAGGGGATGTATCTTATGGTCGATGTGGTTGCTTAACCAT 480
KEYWORDS	GI:41573728	Db	579 TGAGTTCTTCATTGCACGCCAGGGGAATGTACTTCCAGTCCAGCTCGTCGTTAACACG 638
SOURCE		QY	481 TGGGCTATGATGGAGCGGTAGCTCAGTCGATTACAGTGTGTTTAAACCGTTTCAGTTCCC 540
ORGANISM	Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)	Db	639 TCGCTGCCAGCTAGTTCTTACTCTTCCAAACCTCCGAATCTTACGGTCTCTTCAGTACTT 698
	Cryptococcus neoformans var. neoformans	QY	541 AAGACTACTTTCACCCGTTCTGTTTTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGG 600
	Eukaryota; Fungi; Basidiomycota; Hymenomycetes;	Db	699 CTGAAGACTTACCAACCTTTCTGCTGGATCACTTGACTACAAACCAACAGACCAACGTTGAGC 758
	Filobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;	QY	601 ATTGTGCTAGGAGATAACACTGTCTCTCTGCTGATCTCGATACCAACCAAGGATGTGG 660
	Filobasidiella.	Db	759 AATGTTGGTTGGTAGCGACTCTGTGCTTTGGCCGACTTGAACACTGAGAGTGACACTG 818
REFERENCE	1 (bases 1 to 865)	QY	661 TCAAGAATCAATGGTACGACTGGGTGGGATCAATGGTATCGAACTAC 707
AUTHORS	Loftus,B.	Db	819 TTGTTTCTTACTGGGACCAATGGATCACAGACCTTGTGTTTCCAACTAC 865

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:29:19 ; Search time 1.08108 Seconds
(without alignments)
483.353 Million cell updates/sec

Title: us-10-820-200-2_copy_161_167

Perfect score: 40

Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A-COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B-COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A-COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B-COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCITUS-COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	478	1	US-08-720-899-7
2	40	100.0	478	1	US-08-459-610-7
3	40	100.0	478	2	US-08-343-804-7
4	40	100.0	478	2	US-08-339-715A-2
5	40	100.0	478	2	US-08-600-908A-10
6	40	100.0	478	3	US-08-683-838A-10
7	40	100.0	478	3	US-09-182-859-7
8	40	100.0	478	4	US-09-672-459-7
9	40	100.0	478	4	US-09-636-252A-10
10	40	100.0	478	4	US-10-186-042-7
11	33	82.5	180	4	US-09-248-796A-22820
12	32	80.0	84	4	US-09-583-110-4016
13	32	80.0	95	4	US-09-107-433-4442
14	32	80.0	369	1	US-08-700-359-21
15	32	80.0	583	2	US-08-616-392C-4
16	32	80.0	788	4	US-09-107-532A-6538
17	32	80.0	792	4	US-09-134-000C-5895
18	32	80.0	820	2	US-08-380-182-23
19	32	80.0	1118	4	US-09-949-016-6596
20	32	80.0	1120	4	US-09-949-016-10404
21	31	77.5	153	4	US-09-328-352-4170
22	31	77.5	279	4	US-09-270-767-34772
23	31	77.5	279	4	US-09-270-767-49989
24	31	77.5	461	4	US-09-198-452A-924
25	31	77.5	472	4	US-09-438-185A-858
26	31	77.5	474	4	US-09-081-385-148
27	31	77.5	2697	4	US-10-144-198-12

28	30	75.0	25	1	US-08-485-455D-41	Sequence 41, Appl
29	30	75.0	25	2	US-08-482-130C-41	Sequence 41, Appl
30	30	75.0	25	2	US-08-484-211C-41	Sequence 41, Appl
31	30	75.0	25	3	US-08-906-769-41	Sequence 41, Appl
32	30	75.0	25	3	US-08-906-616-41	Sequence 41, Appl
33	30	75.0	25	3	US-08-817-795-41	Sequence 41, Appl
34	30	75.0	25	3	US-08-485-443B-41	Sequence 41, Appl
35	30	75.0	25	3	US-08-639-075A-41	Sequence 41, Appl
36	30	75.0	25	3	US-09-012-431-41	Sequence 41, Appl
37	30	75.0	25	3	US-09-012-692-41	Sequence 41, Appl
38	30	75.0	25	3	US-08-906-613-41	Sequence 41, Appl
39	30	75.0	25	5	PCT-US95-14442A-41	Sequence 41, Appl
40	30	75.0	64	1	US-08-485-455D-75	Sequence 75, Appl
41	30	75.0	64	2	US-08-482-130C-75	Sequence 75, Appl
42	30	75.0	64	2	US-08-484-211C-75	Sequence 75, Appl
43	30	75.0	64	3	US-08-906-769-75	Sequence 75, Appl
44	30	75.0	64	3	US-08-906-616-75	Sequence 75, Appl
45	30	75.0	64	3	US-08-817-795-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1

US-08-720-899-7

; Sequence 7, Application US/08720899

; Patent No. 5753460

; GENERAL INFORMATION:

; APPLICANT: Bisgaard-Frantzen, Henrik

; APPLICANT: Borchert, Torben Vedel

; APPLICANT: Svendsen, Allan

; APPLICANT: Thellersen, Marianne

; APPLICANT: Van der Zee, Pia

; TITLE OF INVENTION: AMYLASE VARIANTS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.

; STREET: 405 Lexington Avenue, 64th Floor

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/720,899

; FILING DATE: 10-OCT-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/343,804

; FILING DATE: 22-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lowney Dr., Karen A.

; REGISTRATION NUMBER: 31,274

; REFERENCE/DOCKET NUMBER: 4054.214-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123

; TELEFAX: 212-878-9655

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 478 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-720-899-7

Query Match 100.0%; Score 40; DB 1; Length 478;

Best Local Similarity 100.0%; Pred No. 4,7; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 SSQDYFH 7
Db      141 SSQDYFH 147

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043o No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match      100.0%; Score 40; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSQDYFH 7
Db      141 SSQDYFH 147

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: No. 5830837o No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-343-804-7

Query Match      100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSQDYFH 7
Db      141 SSQDYFH 147

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; TITLE OF INVENTION: NEOPULLULANASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306096/1993

```



```

; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Melsner, Allen S.
; REGISTRATION NUMBER: 27,215
; REFERENCE/DOCKET NUMBER: 18335.009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 SSQDYFH 147

RESULT 5
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5989169o No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-600-908A-10

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 SSQDYFH 147

RESULT 6
US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6022724o No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,838A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-683-838A-10

Query Match 100.0%; Score 40; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 SSQDYFH 147

RESULT 7
US-09-182-859-7
; Sequence 7, Application US/09182859
; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgaard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
```

```
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-182-853-7
```

```
Query Match          100.0%; Score 40; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSQDYFH 7
Db      141 SSQDYFH 147
```

```
RESULT 8
US-09-672-459-7
; Sequence 7, Application US/09672459
; Patent No. 6436888
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/672,459
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-672-459-7
```

```
Query Match          100.0%; Score 40; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSQDYFH 7
Db      141 SSQDYFH 147
```

```
RESULT 9
US-09-636-252A-10
; Sequence 10, Application US/09636252A
; Patent No. 6440716
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/1F216-US2
; CURRENT APPLICATION NUMBER: US/09/636,252A
```

```
; CURRENT FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683,838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 478
; TYPE: PRT
; ORGANISM: A. oryzae
US-09-636-252A-10
```

```
Query Match          100.0%; Score 40; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSQDYFH 7
Db      141 SSQDYFH 147
```

```
RESULT 10
US-10-186-042-7
; Sequence 7, Application US/10186042
; Patent No. 6642044
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/186,042
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/672,459
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-186-042-7
```

```
Query Match          100.0%; Score 40; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSQDYFH 7
Db      141 SSQDYFH 147
```

```
RESULT 11
US-09-248-796A-22820
; Sequence 22820, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
```

```

/ SOFTWARE: <Unknown>
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,433
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/ 085131
/ FILING DATE: May 12, 1998
/ APPLICATION NUMBER: 60/051553
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: STC-011
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 4442:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 95 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Streptococcus pneumoniae
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (B) LOCATION 1...95
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4442:
/
/ US-09-107-433-4442
/
/ Query Match 80.0%; Score 32; DB 4; Length 95;
/ Best Local Similarity 100.0%; Pred. No. 32;
/ Matches 5; Conservative 0; Mismatches 0; Indels 0;
/
/ Qy 3 QDYFH 7
/ Db 51 QDYFH 55
/
/ RESULT 14
/ US-08-700-359-21
/ Sequence 21, Application US/08700359
/ Patent No. 5766925
/ GENERAL INFORMATION:
/ APPLICANT: SUGIMOTO, MASAKAZU
/ APPLICANT: USUDA, YOSHIIRO
/ APPLICANT: SUZUKI, TOMOKO
/ APPLICANT: TANAKA, AKIKO
/ APPLICANT: MATSUI, HIROSHI
/ TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
/ STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
/ CITY: ARLINGTON
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: MS-DOS TEXT EDITOR
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/700,359
/ FILING DATE: 08-OCT-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP-35019
/ FILING DATE: 04-MAR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: NORMAN F. OBLON

```

REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-819-0 PCT
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-700-359-21

Query Match 80.0%; Score 32; DB 1; Length 369;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 133 SSMYYH 139

RESULT 15

US-08-616-392C-4
Sequence 4, Application US/08616392C
Patent No. 5998165
GENERAL INFORMATION:
APPLICANT: Goold, Richard D.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Sellhauer, Jeffrey
APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCIA
TITLE OF INVENTION: AND PANC1B ASSOCIATED WITH PANCREATIC CANCER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,392C
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/581,240
FILING DATE: 29-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0052-1US
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: hnt
CLONE: 496071
US-08-616-392C-4

Query Match 80.0%; Score 32; DB 2; Length 583;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 433 STDDYFH 439

Search completed: February 8, 2005, 18:44:38
Job time : 2.74775 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:42:07 ; Search time 3.01351 Seconds
(without alignments)
756.564 Million cell updates/sec

Title: US-10-820-200-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	478	14	US-10-184-771-10
2	40	100.0	478	14	US-10-186-042-7
3	40	100.0	478	15	US-10-644-187-7
4	40	100.0	478	17	US-10-926-720-10
5	40	100.0	498	16	US-10-815-495-22
6	40	100.0	499	16	US-10-815-495-18
7	40	100.0	1095	14	US-10-228-063-45
8	34	85.0	256	14	US-10-298-638-16
9	33	82.5	133	15	US-10-424-599-175172
10	33	82.5	155	15	US-10-424-599-218859
11	33	82.5	167	16	US-10-767-701-39199
12	33	82.5	193	16	US-10-437-963-140941
13	33	82.5	703	15	US-10-425-114-65668

14	33	82.5	4351	10	US-09-970-944-20	Sequence 20, Appl
15	32	80.0	71	15	US-10-424-599-162174	Sequence 162174, Ap
16	32	80.0	119	15	US-10-264-049-4124	Sequence 4124, Ap
17	32	80.0	180	17	US-10-472-928-496	Sequence 496, App
18	32	80.0	323	15	US-10-282-122A-72639	Sequence 72639, A
19	32	80.0	583	9	US-09-924-654-4	Sequence 4, Appli
20	32	80.0	594	9	US-09-925-300-1079	Sequence 1079, Ap
21	32	80.0	820	15	US-10-369-493-686	Sequence 686, App
22	32	80.0	820	15	US-10-282-122A-42805	Sequence 42805, A
23	32	80.0	820	15	US-10-282-122A-74838	Sequence 74838, A
24	32	80.0	820	15	US-10-282-122A-75391	Sequence 75391, A
25	32	80.0	820	15	US-10-272-107-3	Sequence 3, Appli
26	32	80.0	820	16	US-10-689-395-3	Sequence 3, Appli
27	32	80.0	885	16	US-10-437-963-136212	Sequence 136212, Ap
28	32	80.0	1118	14	US-10-153-668-104	Sequence 104, App
29	31	77.5	37	17	US-10-472-928-1866	Sequence 1866, Ap
30	31	77.5	69	9	US-09-984-245-184	Sequence 184, App
31	31	77.5	69	10	US-09-966-262-184	Sequence 184, App
32	31	77.5	69	10	US-09-983-966-184	Sequence 184, App
33	31	77.5	69	14	US-10-059-395-184	Sequence 184, App
34	31	77.5	69	14	US-10-143-090-184	Sequence 184, App
35	31	77.5	104	9	US-09-814-661A-2	Sequence 2, Appli
36	31	77.5	191	9	US-09-796-692-1281	Sequence 1281, Ap
37	31	77.5	191	9	US-09-796-692-1682	Sequence 1682, Ap
38	31	77.5	191	9	US-09-796-692-1754	Sequence 1754, Ap
39	31	77.5	191	9	US-09-796-692-1998	Sequence 1998, Ap
40	31	77.5	191	9	US-09-796-692-2036	Sequence 2036, Ap
41	31	77.5	191	14	US-10-040-862-1281	Sequence 1281, Ap
42	31	77.5	191	14	US-10-040-862-1682	Sequence 1682, Ap
43	31	77.5	191	14	US-10-040-862-1754	Sequence 1754, Ap
44	31	77.5	191	14	US-10-040-862-1998	Sequence 1998, Ap
45	31	77.5	191	14	US-10-040-862-2036	Sequence 2036, Ap

ALIGNMENTS

RESULT 1
US-10-184-771-10
; Sequence 10, Application US/10184771
; Publication No. US20030170769A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/1F216-US2
; CURRENT APPLICATION NUMBER: US/10/184,771
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/636,252
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683,838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 478
; TYPE: PRT
; ORGANISM: A. oryzae
US-10-184-771-10

Query Match 100.0%; Score 40; DB 14; Length 478;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7

Db 141 SSQDYFH 147

RESULT 2

US-10-186-042-7

; Sequence 7, Application US/10186042

Publication No. US20030171236A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/186,042
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/672,459
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-186-042-7

Query Match 100.0%; Score 40; DB 14; Length 478;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 SSQDYFH 147

RESULT 3
US-10-644-187-7
; Sequence 7, Application US/10644187
; Publication No. US20040048351A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/644,187
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-644-187-7

Query Match 100.0%; Score 40; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7

DB 141 SSQDYFH 147

RESULT 4

US-10-926-720-10
; Sequence 10, Application US/10926720
; Publication No. US20050019886A1
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: -Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/926,720
; FILING DATE: 26-Aug-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-926-720-10

Query Match 100.0%; Score 40; DB 17; Length 478;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 SSQDYFH 147

RESULT 5

US-10-815-495-22
; Sequence 22, Application US/10815495
; Publication No. US20040191864A1
; GENERAL INFORMATION:
; APPLICANT: Novozymes Biotech, Inc.
; APPLICANT: Connelly, Mariah
; APPLICANT: Brody, Howard
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient
; TITLE OF INVENTION: Mutants Of Aspergillus Niger
; FILE REFERENCE: 10345.200-US
; CURRENT APPLICATION NUMBER: US/10/815,495
; CURRENT FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-815-495-22

Query Match 100.0%; Score 40; DB 16; Length 498;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
|||||||
Db 162 SSQDYFH 168

RESULT 6
US-10-815-495-18
; Sequence 18, Application US/10815495
; Publication No. US20040191864A1
; GENERAL INFORMATION:
; APPLICANT: Novozymes Biotech, Inc.
; APPLICANT: Connelly, Mariah
; APPLICANT: Brody, Howard
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient
; TITLE OF INVENTION: Mutants Of Aspergillus Niger
; FILE REFERENCE: 10345.200-US
; CURRENT APPLICATION NUMBER: US/10/815,495
; CURRENT FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-815-495-18

Query Match 100.0%; Score 40; DB 16; Length 499;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
|||||||
Db 162 SSQDYFH 168

RESULT 7
US-10-228-063-45
; Sequence 45, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Aspergillus shirousami
US-10-228-063-45

Query Match 100.0%; Score 40; DB 14; Length 1095;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
|||||||
Db 141 SSQDYFH 147

RESULT 8

US-10-298-638-16
; Sequence 16, Application US/10298638
; Publication No. US20030177527A1
; GENERAL INFORMATION:
; APPLICANT: HARING, Michel A. et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE
; TITLE OF INVENTION: STRATEGIES BASED ON PLANT DEFENCE
; FILE REFERENCE: 2121-0167P
; CURRENT APPLICATION NUMBER: US/10/298,638
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 10/298,638
; PRIOR FILING DATE: 2002-11-19
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Arabidopsis arenosa
US-10-298-638-16

Query Match 85.0%; Score 34; DB 14; Length 256;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
|||||||
Db 164 SSQDYFH 170

RESULT 9

US-10-424-599-175172
; Sequence 175172, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175172
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129199C.1.pep
US-10-424-599-175172

Query Match 82.5%; Score 33; DB 15; Length 133;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SODYFH 7
|:|||||
Db 98 SEDYFH 103

RESULT 10

US-10-424-599-218859
; Sequence 218859, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218859
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39658C.1.pep
US-10-424-599-218859

Query Match 82.5%; Score 33; DB 15; Length 155;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|:|||||
Db 7 SARDYFH 13

RESULT 11
US-10-767-701-39199
; Sequence 39199, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39199
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C99243_1.pep
US-10-767-701-39199

Query Match 82.5%; Score 33; DB 16; Length 167;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|:|||||
Db 112 SAKDYFH 118

RESULT 12
US-10-437-963-140941
; Sequence 140941, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140941

; LENGTH: 193
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42092C.1.pep
US-10-437-963-140941

Query Match 82.5%; Score 33; DB 16; Length 193;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|:|||||
Db 138 SAKDYFH 144

RESULT 13
US-10-425-114-65668
; Sequence 65668, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65668
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700333159_FLI.pep
US-10-425-114-65668

Query Match 82.5%; Score 33; DB 15; Length 703;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SQDYFH 7
|:|||||
Db 82 SEDYFH 87

RESULT 14
US-09-970-944-20
; Sequence 20, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shankets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 4351
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-970-944-20

Query Match 82.5%; Score 33; DB 10; Length 4351;
Best Local Similarity 71.4%; Pred. No. 3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
:|||||
Db 605 NEQDYFH 611

RESULT 15
US-10-424-599-162174
; Sequence 162174, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162174
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(71)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117460C.1.pap
US-10-424-599-162174

Query Match 80.0%; Score 32; DB 15; Length 71;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SQDYFH 7
:|||||
Db 21 SKDYFH 26

Search completed: February 8, 2005, 19:05:54
Job time : 4.34685 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:22:38 ; Search time 4.01351 Seconds
(without alignments)
674.552 Million cell updates/sec

Title: US-10-820-200-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	423	4	ABB09072 Aspergill
2	40	100.0	478	2	Aar46065 Mutant al
3	40	100.0	478	2	Aar72450 Aspergill
4	40	100.0	478	2	Aar78270 Aspergill
5	40	100.0	478	2	Aar79025 Mature ta
6	40	100.0	478	2	Aaw14500 Aspergill
7	40	100.0	498	4	Aab84206 Amino aci
8	40	100.0	498	8	Adt89632 Aspergill
9	40	100.0	499	8	Adt89628 Aspergill
10	40	100.0	1095	6	Abp96630 Alpha-amy
11	37	92.5	493	2	Aar88212 Alpha-amy
12	34	85.0	52	4	Aao08447 Human pol
13	34	85.0	1222	3	Aag36453 Arabidops
14	34	85.0	1257	3	Aag36452 Arabidops
15	34	85.0	1275	3	Aag36451 Arabidops
16	33	82.5	116	5	Abp08801 Human ORF
17	33	82.5	124	4	Aao05772 Human pol
18	33	82.5	342	6	Abu070873 Human adl
19	33	82.5	379	8	Ado58298 S. domunc
20	33	82.5	2112	7	Adg76988 Human nuc
21	33	82.5	3124	8	Adj78494 EIA bindi
22	33	82.5	4351	7	Adg42622 Rat MEGFI
23	32	80.0	82	8	Adn05042 Antipeori
24	32	80.0	82	8	Adp25302 PRO polyp
25	32	80.0	84	8	Adk47501 Streptoco

26	32	80.0	95	8	ADR95807	Novel S.
27	32	80.0	119	5	ABP42992	Human ova
28	32	80.0	180	6	ABU00681	S. pneumo
29	32	80.0	180	8	ADM92084	S. pneumon
30	32	80.0	182	5	ABP28341	Streptoco
31	32	80.0	195	8	ADH45430	Human mol
32	32	80.0	223	6	ABU44715	Protein e
33	32	80.0	326	6	ABP98616	Mature Er
34	32	80.0	348	6	ABP98615	Trwinia c
35	32	80.0	351	8	ADO61605	Transcrip
36	32	80.0	501	4	ABG15126	Novel hum
37	32	80.0	529	4	ABR70891	Drosophil
38	32	80.0	583	2	AAW26654	Human PAN
39	32	80.0	583	6	ABG72922	Novel hum
40	32	80.0	583	8	ABM81509	Tumour-as
41	32	80.0	583	8	ADP25358	PRO polyp
42	32	80.0	594	3	AAJ56501	Human pro
43	32	80.0	685	7	ADK62434	Disease t
44	32	80.0	700	2	AAJ33673	B. bassia
45	32	80.0	715	4	AAW79737	Human pro

ALIGNMENTS

RESULT 1

ABB09072	26-JUN-2002	(first entry)
ID	ABB09072	standard; protein; 423 AA.
XX	AC	ABB09072;
XX	XX	
DT	26-JUN-2002	(first entry)
DE	XX	
DE	XX	Aspergillus oryzae TAKA protein (TAA).
XX	XX	
KW	Thermus sp. IM6501;	maltogetic amylase; EC 3.2.1.133; crystallisation;
KW	protein co-ordinate data;	3 dimensional structure.
XX	OS	Aspergillus oryzae.
XX	XX	
PN	KR2001027418-A.	
XX	XX	
PD	06-APR-2001.	
XX	XX	
PF	09-SEP-1999;	99KR-00039130.
XX	XX	
PR	09-SEP-1999;	99KR-00039130.
XX	XX	
PA	(POST-) POSTECH FOUND.	
PA	(SAMV-) SAMYANG GENEX CORP.	
XX	XX	
PI	Kim TJ, Park GH;	
XX	XX	
DR	WPI; 2001-534477/59.	
XX	XX	
PT	Manufacturing maltogetic amylase having improved transglycosylation	
PT	activity, comprises using crystallization.	
XX	XX	
PS	Disclosure; Page 188; 196pp; Korean.	
XX	XX	
CC	XX	
CC	XX	The present invention describes manufacturing maltogetic amylase (EC
CC	XX	3.2.1.133) having improved transglycosylation activity, comprising using
CC	XX	crystallisation and the three dimensional structure of maltogetic
CC	XX	amylase. Manufacturing maltogetic amylase comprises the following steps:
CC	XX	(i) obtaining a gene of maltogetic amylase from Thermus sp. IM6501 (KCTC
CC	XX	5027BP) and inserting the gene into plasmid pUC119 to construct
CC	XX	recombinant DNA (pTHW119); (ii) inserting the recombinant DNA to
CC	XX	Escherichia coli MC1061, which is cultivated at 37 plus degrees Celsius
CC	XX	for 10 hours in Luria-Bertani (LB) media and centrifuging the media to
CC	XX	obtain a microbial cell; (iii) suspending the microbial cell with buffer
CC	XX	solution at pH 7.5 and obtaining supernatant; and (iv) passing the
CC	XX	supernatant through column chromatography and obtaining purified
CC	XX	maltogetic amylase. The maltogetic amylase is a dimer comprised of two

CC maltogenic amylase molecules, and a *Thermus* sp. IM6501 maltogenic amylase (ThMA) crystal. The amylase has a structure containing an activated region that consists of amino acid residues of Asp-328, Glu-357, Asp-424, and a pocket with glucose bound that consists of amino acid residues of CC Pro44, Tyr45, Arg81, Arg83, Pro118, Cys116, Asn131, Glu132, Val1329, CC and His1360. The present sequence represents *Aspergillus oryzae* TAKA CC protein (TAA), given in comparison with ThMA in the present invention XX
SQ Sequence 423 AA;

Query Match 100.0%; Score 40; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
Db 137 SSQDYFH 143

RESULT 2
AAR46065
ID AAR46065 standard; protein; 478 AA.

XX AAR46065;
XX 25-MAR-2003 (revised)
DT 18-JUL-1994 (first entry)
XX
XX Mutant alpha-amylase.

XX Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liquifaction agents.

OS *Aspergillus oryzae*.
XX
XX WO9402597-A1.

XX 03-FEB-1994.
XX
XX 06-JUL-1993; 93WO-DK000230.

XX 23-JUL-1992; 92DK-00000946.
PR 16-DEC-1992; 92DK-00001503.
PR 15-MAR-1993; 93DK-00000292.

XX (NOVO) NOVO-NORDISK AS.

XX Svendsen A, Bisgard-Frantzen H;
XX WPI; 1994-048855/06.

XX Mutant alpha-amylase from *Bacillus* species comprising a methionine substitution - with improved stability and activity at low pH, for use in PT detergents, dishwashing agents and liquifaction agents.

XX Claim 1; Page 7; 20pp; English.

XX The sequence as that of the *Aspergillus oryzae* alpha amylase, sold CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can be CC mutated by substitution of one or more of its methionine residues for any CC amino acid other than cysteine. The mutant alpha-amylase exhibits a CC better activity level and better stability in the presence of oxidising CC agents than previous mutant alpha amylases, and improved thermostability CC at moderately low pH. The enzyme can be used as an additive for CC detergents, dishwashing agents and liquifaction agents. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 478 AA;

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
Db 141 SSQDYFH 147

RESULT 3
AAR72450
ID AAR72450 standard; protein; 478 AA.

XX AAR72450;

XX 25-MAR-2003 (revised)
DT 01-DEC-1995 (first entry)

XX *Aspergillus oryzae* alpha amylase (mature protein).

XX Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; *Bacillus licheniformis*; *Bacillus amyloliquefaciens*;
KW *Bacillus stearothermophilus*; dyeing; bleaching; scouring; textile;
KW thermostable.

OS *Aspergillus oryzae*.

XX WO9510603-A1.

XX 20-APR-1995.

XX 05-OCT-1994; 94WO-DK000370.

XX 08-OCT-1993; 93DK-00001133.

PR 02-FEB-1994; 94DK-00000140.

XX (NOVO) NOVO-NORDISK AS.

XX Borchert TV, Bisgard-Frantzen H, Svendsen A, Thellersen M;
PI Van Der Zee P;

XX WPI; 1995-161790/21.

XX New *Bacillus* derived alpha-amylase variants - having amino acid PT modifications to improve washing and/or dishwashing performance.

XX Disclosure; Page 75-76; 105pp; English.

XX Variant alpha amylase enzymes which have improved washing and/or as CC detergent additives. The enzymes have one or more amino acid residues CC added, deleted or substituted. The variants can also be used for textile CC desizing prior to scouring, bleaching and dyeing. The variants have CC improved thermostability, acid/alkaline stability; low temperature CC optimum; pH optimum; higher hydrolysis velocity and improved tolerance to CC other composition constituents, e.g. oxidation agents. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 478 AA;

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSQDYFH 7
Db 141 SSQDYFH 147

RESULT 4
AAR78270
ID AAR78270 standard; protein; 478 AA.

XX AAR78270;

XX 17-JAN-1996 (first entry)

XX *Aspergillus oryzae* alpha amylase (mature protein).

XX Alpha amylase; oxidation; desizing; bleaching; scouring; fabric; starch;
 KW thermostable; methionine; Bacillus licheniformis;
 KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;
 KW Aspergillus oryzae.
 XX
 XX Aspergillus oryzae.
 XX
 XX WO9521247-A1.
 XX
 XX 10-AUG-1995.
 XX
 XX 05-OCT-1994; 94WO-DK000371.
 XX
 XX 02-FEB-1994; 94DK-00000141.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 XX
 XX Toft AH, Marcher D, Pedersen HH, Nilsson TE;
 XX WPI; 1995-283767/37.
 XX
 XX Use of an oxidation stable alpha-amylase - for simultaneous desizing and
 PT bleaching or scouring of fabrics contg. starch or starch derivs.
 XX
 XX Disclosure; Page 25-26; 37pp; English.
 XX
 XX Oxidation stable alpha amylases can be used for the simultaneous desizing
 CC and bleaching or scouring of a fabric comprising starch or starch
 CC derivatives. They exhibit a better heat stability, especially in the
 CC presence of oxidising agents. They are obtained from a parent alpha
 CC amylase by replacing one or more methionine residues with any amino acid
 CC different from Cys or Met, preferably Leu, Thr, Ala, Gly, Ser, Ile or
 CC Asp. The parent alpha amylase is pref. derived from a Bacillus species,
 CC although alpha amylases of fungal origin can also be used. This sequence
 CC is the wild type (unmodified) alpha amylase of Aspergillus oryzae.
 XX
 XX SQ Sequence 478 AA;
 Query Match 100.0%; Score 40; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSQDYFH 7
 Db 141 SSQDYFH 147
 RESULT 5
 AAR79025
 ID AAR79025 standard; protein; 478 AA.
 XX
 XX AAR79025;
 XX
 XX 22-MAR-1996 (first entry)
 XX
 XX Mature taka-amylase A.
 DE
 DE Wild type; neopullulanase; B. stearothermophilus; mutant; food industry;
 KW modification; hydrophobicity; replacement insertion; deletion.
 XX
 XX Aspergillus oryzae.
 OS
 XX Key Location/Qualifiers
 FH Disulfide-bond 30..38
 FT Disulfide-bond 150..164
 FT Disulfide-bond 240..283
 FT Disulfide-bond 439..474
 XX
 XX JP07177891-A.
 PN
 XX 18-JUL-1995.
 XX
 XX

PF 31-OCT-1994; 94JP-00288658.
 XX
 PR 12-NOV-1993; 93JP-00306096.
 XX
 XX (NIDE) NEC CORP.
 PA (EZAK) EZAKI GLICO CO.
 XX
 XX WPI; 1995-279919/37.
 XX
 XX Modifying a transferase by enhancing hydrophobicity of a selected site -
 PT increases transfer activity, also new mutant neo-pullulanase(s).
 PT
 PS Disclosure; Page 10-11; 18pp; English.
 XX
 XX This sequence represents the mature form of taka-amylase from A. oryzae.
 CC This sequence was used in a method for the generation of mutant
 CC pullulanases for use in the food industry (see also AAR79026-28). The
 CC wild type pullulanase enzyme was modified by the method of the invention
 CC for enhancing the hydrophobicity of a selected site of the pullulanase.
 CC The method comprises replacement of a group in the selected site with a
 CC hydrophobic group, replacement of an amino acid with a hydrophobic amino
 CC acid, and/or insertion or deletion of a hydrophobic amino acid from the
 CC selected site. The method was used to produce neopullulanases Y377F,
 CC S422V and M375L
 XX
 XX SQ Sequence 478 AA;
 Query Match 100.0%; Score 40; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSQDYFH 7
 Db 141 SSQDYFH 147
 RESULT 6
 AAW14500
 ID AAW14500 standard; protein; 478 AA.
 XX
 XX AAW14500;
 XX
 XX 04-JUN-1997 (first entry)
 DT
 XX Aspergillus oryzae alpha-amylase (mature protein).
 DE
 DE alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;
 KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
 KW calcium dependency; substrate binding; stability; pH optimum;
 KW thermostability; cleavage; oligosaccharide substrate; dishwashing;
 KW washing; detergent additive; fabric desizing; starch liquefaction;
 KW sweetener; ethanol production; variant.
 KW
 XX Aspergillus oryzae.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 13..45
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 7-23 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 33"
 FT Misc-difference 14..40
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 8-18 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to this
 FT fragment; claim 35"
 FT Misc-difference 28..42
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding

FT to 12-19 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 30"
 FT Misc-difference 32. .38
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 14-15 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to this
 FT fragment; claim 32"
 FT Misc-difference 66. .84
 FT /label= loop 2 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 44-57 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 18"
 FT Misc-difference 70. .78
 FT /label= loop 2 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 48-51 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to this
 FT fragment; claim 20"
 FT Misc-difference 98. .210
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 117-185 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 24"
 FT Misc-difference 102. .206
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 121-181 of AAW14499 is
 FT deleted or replaced with a fragment corresponding to this
 FT fragment; claim 26"
 FT Misc-difference 121. .181
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to this fragment is deleted or replaced with a fragment
 FT corresponding to 102-206 of AAW14499; claim 41"
 FT Misc-difference 121. .174
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to this fragment is deleted or
 FT replaced with a fragment corresponding to 102-199 of
 FT AAW14499; claim 42"
 FT Misc-difference 165. .177
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 195-202 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 21"
 FT Misc-difference 166. .173
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for
 FT a variant) corresponding to 196-198 of AAW14499 is
 FT deleted or replaced with a fragment corresponding to this
 FT fragment; claim 23"
 FT Misc-difference 181. .184
 FT /note= "an amino acid fragment corresponding to this
 FT region is deleted from the parent sequence of a variant
 FT Fungamyl; claim 43"
 FT Misc-difference 291. .313
 FT /label= loop 8 modification region
 FT /note= "at least one amino acid residue of a parent alpha
 FT -amylase (used as a template for a variant) corresponding
 FT to 322-346 of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim 36"
 FT Misc-difference 297. .313
 FT /label= loop 8 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a template for

FT a variant) corresponding to 325-345 of AAW14498 is
 FT deleted or replaced with a fragment corresponding to this
 FT fragment; claim 38"
 XX WO9623874-A1.
 XX 08-AUG-1996.
 XX 05-FEB-1996; 96WO-DK000057.
 XX 03-FEB-1995; 95DK-00000128.
 XX 23-OCT-1995; 95DK-00001192.
 XX 10-NOV-1995; 95DK-00001256.
 XX (NOVO) NOVO-NORDISK AS.
 XX Svendsen A, Bisgard-Frantzen H, Borchert TV;
 XX WPI; 1996-371424/37.
 XX Alpha-amylase variants and methods of production - have altered
 XX PT properties such as calcium dependency, substrate binding and stability.
 XX PS Disclosure; Page 87-88; 171pp; English.
 XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
 CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
 CC methods of constructing them) are claimed. Examples of variants are
 CC featured above. The variants have altered properties such as calcium
 CC dependency, substrate binding and stability. Also one or more proline or
 CC cysteine residues in the variant is modified or replaced with a non-
 CC proline or non-cysteine residue such as alanine. The variants can be used
 CC for (dish)washing, as detergent additives or for fabric desizing or
 CC starch liquefaction. They can also be used for the production of
 CC sweeteners and ethanol from starch. See also AAW14498-99
 XX SQ Sequence 478 AA;
 XX
 Query Match 100.0%; Score 40; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSQDYFH 7
 Db 141 SSQDYFH 147
 RESULT 7
 AAB84206
 ID AAB84206 standard; protein; 498 AA.
 XX AAB84206;
 AC AAB84206;
 XX 06-AUG-2001 (first entry)
 DT Amino acid sequence of a fungamyl-like alpha-amylase.
 DE Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;
 XX alcohol; starch; dough improver; brewing; starch liquification.
 KW Aspergillus oryzae.
 XX WO200134784-A1.
 XX 17-MAY-2001.
 PD 10-NOV-2000; 2000WO-DK000626.
 PF 10-NOV-1999; 99DK-00001617.
 XX (NOVO) NOVOZYMES AS.
 PA Bisgard-Frantzen H, Svendsen A, Pedersen S;
 XX

XX WPI; 2001-367478/38.
DR N-PSDB; AAF90208.
XX
XX New variant of Fungamyl-like alpha-amylase, useful for production of
PT maltose syrups, includes mutations that improve stability against heat
PT and acidic pH.
XX
XX Claim 1; Page 42-45; 49pp; English.
XX
XX The present sequence represents a fungamyl-like alpha-amylase. The
CC specification describes variants of this fungamyl-like alpha-amylase, which
CC have an alteration in one the amino acid regions 98-110, 150-160,
CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion or
CC substitution of an amino acid or an insertion of an amino acid downstream
CC of a particular position. The variants retain alpha-amylase activity, and
CC have better heat stability and/or stability at acidic pH, relative to
CC wild-type enzyme. The variants can therefore be used at higher
CC temperatures (more efficient conversion or faster reaction, and have
CC reduced need for cooling and reduced risk of contamination). The variants
CC may also be used in conjunction with other enzymes, particularly
CC Glucoamylase during dextrinisation. The variants are used to produce
CC syrups, particularly of high maltose content, or alcohol, from starch; as
CC dough improver for baked goods; in brewing, to increase fermentability of
CC the wort; and for liquefaction of starch
XX
XX Sequence 498 AA;
SQ

Query Match 100.0%; Score 40; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSQDYFH 7
DB 161 SSQDYFH 167
|||||

RESULT 8
ADT89632
ID ADT89632 standard; protein; 498 AA.
XX
XX AC ADT89632;
XX
XX DT 16-DEC-2004 (first entry)
XX
XX DE Aspergillus niger neutral alpha-amylase A (amyA) protein.
XX KW Glucoamylase; glaA; amyA; alpha-amylase A; enzyme.
XX OS Aspergillus niger.
XX PN US2004191864-A1.
XX PD 30-SEP-2004.
XX PF 31-MAR-2004; 2004US-00815495.
XX PR 31-MAR-2003; 2003US-0459902P.
XX PA (NOVO) NOVOZYMES BIOTECH INC.
XX PI Connelly M, Brody H;
XX DR WPI; 2004-708545/69.
XX DR N-PSDB; ADT89631.
XX
XX PS Producing heterologous biological substance comprises culturing mutant of
PT wild-type Aspergillus niger strain in medium suitable for producing
PT heterologous biological substance and recovering heterologous biological
PT substance.
XX
XX Example 11; SEQ ID NO 22; 58pp; English.
XX

Query Match 100.0%; Score 40; DB 8; Length 499;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSQDYFH 7
DB 162 SSQDYFH 168
|||||

RESULT 9
ADT89628
ID ADT89628 standard; protein; 499 AA.
XX
XX AC ADT89628;
XX
XX DT 16-DEC-2004 (first entry)
XX
XX DE Aspergillus niger neutral alpha-amylase B (amyB) protein.
XX KW Glucoamylase; glaA; amyB; alpha-amylase B; enzyme.
XX OS Aspergillus niger.
XX PN US2004191864-A1.
XX PD 30-SEP-2004.
XX PF 31-MAR-2004; 2004US-00815495.
XX PR 31-MAR-2003; 2003US-0459902P.
XX PA (NOVO) NOVOZYMES BIOTECH INC.
XX PI Connelly M, Brody H;
XX DR WPI; 2004-708545/69.
XX DR N-PSDB; ADT89627.
XX
XX PS Producing heterologous biological substance comprises culturing mutant of
PT wild-type Aspergillus niger strain in medium suitable for producing
PT heterologous biological substance and recovering heterologous biological
PT substance.
XX
XX Example 10; SEQ ID NO 18; 58pp; English.
XX

Query Match 100.0%; Score 40; DB 8; Length 499;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSQDYFH 7
DB 162 SSQDYFH 168
|||||

RESULT 9
ADT89628
ID ADT89628 standard; protein; 499 AA.
XX
XX AC ADT89628;
XX
XX DT 16-DEC-2004 (first entry)
XX
XX DE Aspergillus niger neutral alpha-amylase B (amyB) protein.
XX KW Glucoamylase; glaA; amyB; alpha-amylase B; enzyme.
XX OS Aspergillus niger.
XX PN US2004191864-A1.
XX PD 30-SEP-2004.
XX PF 31-MAR-2004; 2004US-00815495.
XX PR 31-MAR-2003; 2003US-0459902P.
XX PA (NOVO) NOVOZYMES BIOTECH INC.
XX PI Connelly M, Brody H;
XX DR WPI; 2004-708545/69.
XX DR N-PSDB; ADT89627.
XX
XX PS Producing heterologous biological substance comprises culturing mutant of
PT wild-type Aspergillus niger strain in medium suitable for producing
PT heterologous biological substance and recovering heterologous biological
PT substance.
XX
XX Example 10; SEQ ID NO 18; 58pp; English.
XX

```

Db      162 SSQDYFH 168

RESULT 10
ABP96630
ID      ABP96630 standard; protein; 1095 AA.
XX
AC      ABP96630;
XX
XX      02-JUN-2003 (first entry)
XX
XX      Alpha-amylase/glucoamylase fusion protein sequence SEQ ID NO:45.
DE
XX
XX      Self-processing plant; plant; processing enzyme; alpha-amylase; grain;
KW      pullulanase; alpha-glucosidase; glucose isomerase; glucoamylase;
KW      mesophilic; thermophilic; hyperthermophilic; transgenic plant; starch;
KW      maltodextrin; ethanol; fermentation; beverage; enzyme.
XX
XX      Aspergillus shirousami.
OS
XX      Synthetic.
XX
XX      WO2003018766-A2.
PN
XX
XX      06-MAR-2003.
PD
XX
XX      27-AUG-2002; 2002WO-US027129.
PF
XX
XX      27-AUG-2001; 2001US-0315281P.
PR
XX
XX      (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA
XX
XX      Lanahan MB, Basu SS, Batie CJ, Chen W, Craig J, Kinkema M;
PI
XX
XX      WPI; 2003-268420/26.
DR
XX      N-PSDB; ACC44572.
DR
XX
XX      Novel polynucleotide encoding hyperthermophilic processing enzymes e.g.
PT      alpha-amylase, useful for producing plant to produce food products having
PT      improved taste or fermentable substrates for ethanol.
XX
XX      Claim 1; Page 107; 158pp; English.
PS
XX
XX      The present invention describes polynucleotides which encode processing
CC      enzymes (e.g. alpha-amylase, pullulanase, alpha-glucosidase, glucose
CC      isomerase, or glucoamylase) that are optimised for expression in plants.
CC      The polynucleotides encode mesophilic, thermophilic or hyperthermophilic
CC      processing enzymes, which are activated under suitable conditions to act
CC      upon the desired substrate. Also described are self-processing transgenic
CC      plants and plant parts, e.g. grain, which express one or more of these
CC      enzymes and have an altered composition that facilitates plant and grain
CC      processing. Also described is a method (M) for converting starch to
CC      starch-derived products in a transformed plant part (TPP), by activating
CC      the starch processing enzyme contained in it. Transgenic grain is useful
CC      for preparing maltodextrin. A transformed plant (TP) can be used to
CC      produce food products having improved taste and to produce fermentable
CC      substrates for ethanol and fermented beverages. (M) eliminates the need
CC      to mill or physically disrupt the integrity of plant parts prior to
CC      recovery of starch-derived products. The present sequence represents
CC      alpha-amylase/glucoamylase fusion protein, which is given in the
CC      exemplification of the present invention
XX
XX      Sequence 1095 AA;
SQ

Query Match      100.0%; Score 40; DB 6; Length 1095;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSQDYFH 7
      :|||||
Db      141 SSQDYFH 147

RESULT 11
Query Match      100.0%; Score 40; DB 6; Length 1095;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSQDYFH 7
      :|||||
Db      141 SSQDYFH 147

RESULT 12
AAO08447
ID      AAO08447 standard; protein; 52 AA.
XX
XX      AAO08447;
AC
XX
XX      06-NOV-2001 (first entry)
DT
XX
XX      Human polypeptide SEQ ID NO 22339.
DE
XX
XX      Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW      vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW      tissue growth factor; immunomodulatory; cancer; leukaemia;
KW      nervous system disorders; arthritis; inflammation.
XX
XX      Homo sapiens.
OS

```

```

AAR8212
ID      AAR8212 standard; protein; 493 AA.
XX
XX      AAR8212;
AC
XX      16-OCT-2003 (revised)
DT      03-APR-1996 (first entry)
DT
XX      Alpha-amylase.
DE
XX
XX      Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
KW
XX
XX      Thermomyces lanuginosus; CBS 224.63.
OS
XX
XX      Key      Location/Qualifiers
FT      Peptide 1..18
FT      /label= Sig_peptide
XX
XX      WO9601323-A1.
PN
XX
XX      18-JAN-1996.
PD
XX
XX      03-JUL-1995; 95WO-EP002607.
PF
XX
XX      04-JUL-1994; 94GB-00013419.
PR
XX
XX      (DANI-) DANISCO AS.
PA
XX
XX      Michelsen B, Rasmussen P;
PI
XX
XX      WPI; 1996-087673/09.
DR
XX      N-PSDB; AAT10562.
DR
XX
XX      Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT      derived from Thermomyces lanuginosus, useful in the prepn. of foodstuffs
PT      and bakery prods. esp. bread.
XX
XX      Claim 3; Page 36-38; 94pp; English.
PS
XX
XX      A thermostable alpha-amylase (AAR8212) of Thermomyces lanuginosus CBS
CC      224.63 is expressed in Aspergillus niger hosts using a gene (AAT10562)
CC      isolated from a T. lanuginosus gene library. The recombinant enzyme (54-
CC      60 kDa) shows optimal activity at 60-70 deg and pH 5.8-6, has a pI of 3.7
CC      and is active at 60-80 deg. (Updated on 16-OCT-2003 to standardise OS
CC      field)
XX
XX      Sequence 493 AA;
SQ

Query Match      92.5%; Score 37; DB 2; Length 493;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSQDYFH 7
      :|||||
Db      161 NSQDYFH 167

RESULT 12
AAO08447
ID      AAO08447 standard; protein; 52 AA.
XX
XX      AAO08447;
AC
XX
XX      06-NOV-2001 (first entry)
DT
XX
XX      Human polypeptide SEQ ID NO 22339.
DE
XX
XX      Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW      vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW      tissue growth factor; immunomodulatory; cancer; leukaemia;
KW      nervous system disorders; arthritis; inflammation.
XX
XX      Homo sapiens.
OS

```


XX WO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004927.
PF
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX (HYSE-) HYSQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI88378.
DR
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 22339; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: the sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 52 AA;
SQ

Query Match 85.0%; Score 34; DB 4; Length 52;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SSQDYFH 7
DB 2 SSQDYMH 8
|||||
|

RESULT 13
AAG36453
ID AAG36453 standard; protein; 1222 AA.
XX
XX AAG36453;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 44675.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR
XX 05-MAR-1999; 99US-0123180P.
PR
XX 09-MAR-1999; 99US-0123548P.
PR
XX 23-MAR-1999; 99US-0125788P.
PR

25-MAR-1999; 99US-0126264P.
29-MAR-1999; 99US-0126785P.
01-APR-1999; 99US-0127462P.
06-APR-1999; 99US-0128234P.
08-APR-1999; 99US-0128714P.
16-APR-1999; 99US-0129845P.
19-APR-1999; 99US-0130077P.
21-APR-1999; 99US-0130449P.
23-APR-1999; 99US-0130510P.
23-APR-1999; 99US-0130891P.
28-APR-1999; 99US-0131449P.
30-APR-1999; 99US-0132048P.
30-APR-1999; 99US-0132407P.
04-MAY-1999; 99US-0132484P.
05-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
06-MAY-1999; 99US-0132487P.
07-MAY-1999; 99US-0132863P.
11-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
02-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.

```
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161408P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 85.0%; Score 34; DB 3; Length 1222;
Best Local Similarity 85.7%; Pred. No. 7, 5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
Db 1130 SSADYFH 1136

RESULT 14
AAG36452
ID AAG36452 standard; protein; 1257 AA.
AC AAG36452;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44674.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 25-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
```

PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142877P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 23-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.

PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160802P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 85.0%; Score 34; DB 3; Length 1257;
Best Local Similarity 85.7%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 1165 SSADYFH 1171

RESULT 15
AAG36451
ID AAG36451 standard; protein; 1275 AA.

XX AAG36451;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44673.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.

XX EP1033405-A2.
PN
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149303P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.

PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 85.0%; Score 34; DB 3; Length 1275;

Best Local Similarity 85.7%; Pred. No. 7.9e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYFH 7

Db 1183 SSADYFH 1189

Search completed: February 8, 2005, 18:36:42
Job time : 7.34685 secs

THIS PAGE BLANK (usprto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:22:38 ; Search time 285.533 Seconds
(without alignments)
674.552 Million cell updates/sec

Title: US-10-820-200-2

Perfect score: 2684

Sequence: 1 MWAWSLFLYGLQVAAPALA.....LPRVLPTEKLAGSKICSSS 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003Bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2684	100.0	498	4 AAB84206	Aab84206 Amino aci
2	2674	99.6	499	8 ADT89628	Adt89628 Aspergill
3	2662	99.2	498	8 ADT89632	Adt89632 Aspergill
4	2572	95.8	478	2 AAR72450	Aar72450 Aspergill
5	2572	95.8	478	2 AAR78270	Aar78270 Aspergill
6	2572	95.8	478	2 AAW14500	Aaw14500 Aspergill
7	2569	95.7	1095	6 ABP96630	Abp96630 Alpha-amy
8	2552	95.1	478	2 AAR46065	Aar46065 Mutant al
9	2503.5	93.3	478	2 AAR79025	Aar79025 Mature ta
10	2151	80.1	423	4 ABB09072	Abb09072 Aspergill
11	2100	78.2	495	6 ABB80177	Abb80177 A. fumiga
12	2032	75.0	494	6 ABP97894	Abp97894 Amino aci
13	1835	68.4	630	6 ABB80178	Abb80178 A. fumiga
14	1778	66.2	484	5 AAE24207	Aae24207 Aspergill
15	1778	66.2	484	8 ADS75939	Ads75939 Aspergill
16	1666.5	62.1	493	2 AAR88212	Aar88212 Alpha-amy
17	1449	54.0	512	2 AAR07574	Aar07574 Alpha-amy
18	1400.5	52.2	511	1 AAP81180	Aap81180 Sequence
19	1398.5	52.1	511	1 AAP81161	Aap81161 Recombina
20	1397	52.0	468	2 AAR24136	Aar24136 Alpha-amy
21	1396.5	52.0	494	1 AAP70571	Aap70571 Alpha-amy
22	1392	51.9	468	2 AAR63184	Aar63184 Variant a
23	1391	51.8	468	2 AAR63185	Aar63185 Variant a
24	1387	51.7	468	2 AAR63187	Aar63187 Variant a
25	1338	49.9	452	2 AAR63186	Aar63186 Variant a

26	1214.5	45.2	547	6 ABB80179	Abb80179 A. fumiga
27	1183	44.1	555	6 ABP97896	Abp97896 Amino aci
28	1163	43.3	567	6 ABP97899	Abp97899 Amino aci
29	1160	43.2	500	8 ADS23436	Ads23436 Bacterial
30	1085	40.4	549	6 ABP97895	Abp97895 Amino aci
31	1015.5	37.8	513	8 ADN19570	Adn19570 Bacterial
32	898.5	33.5	564	8 ADS44249	Ads44249 Bacterial
33	873	32.5	493	8 ADS44293	Ads44293 Bacterial
34	541	20.2	502	6 ABU03092	Abu03092 Alpha amy
35	483.5	18.0	547	8 ADS30907	Ads30907 Bacterial
36	480.5	17.9	547	8 ADS30029	Ads30029 Bacterial
37	456.5	17.0	713	7 ADC23483	Adc23483 Bacillus
38	453	16.9	719	4 AAB74220	Aab74220 Bacillus
39	452.5	16.9	712	2 AAR10051	Aar10051 Cyclonalt
40	452	16.8	719	4 AAB74219	Aab74219 Bacillus
41	450	16.8	719	2 AAY31731	Aay31731 Bacillus
42	450	16.8	719	2 AAY30621	Aay30621 Amino aci
43	450	16.8	719	3 AAY94271	Aay94271 Bacillus
44	450	16.8	719	4 AAB59956	Aab59956 Bacillus
45	450	16.8	719	4 AAB74216	Aab74216 Bacillus

ALIGNMENTS

RESULT 1

AAB84206

ID AAB84206 standard; protein; 498 AA.

AC AAB84206;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of a fungamyl-like alpha-amylase.

KW Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose; alcohol; starch; dough improver; brewing; starch liquification.

OS Aspergillus oryzae.

PN WO200134784-A1.

PD 17-MAY-2001.

PF 10-NOV-2000; 2000WO-DK000626.

PR 10-NOV-1999; 99DK-00001617.

PA (NOVO) NOVOZYMES AS.

PI Bisgard-Frantzen H, Svendsen A, Pedersen S;

DR WPI: 2001-367478/38.

DR N-PSDB; AAF90208.

XX New variant of Fungamyl-like alpha-amylase, useful for production of maltose syrups, includes mutations that improve stability against heat and acidic pH.

PS Claim 1; Page 42-45; 49pp; English.

XX The present sequence represents a fungamyl-like alpha-amylase. The specification describes variants of this fungamyl-like alpha-amylase, which have an alteration in one the amino acid regions 98-110, 150-160, 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion or substitution of an amino acid or an insertion of an amino acid downstream of a particular position. The variants retain alpha-amylase activity, and have better heat stability and/or stability at acidic pH, relative to wild-type enzyme. The variants can therefore be used at higher temperatures (more efficient conversion or faster reaction, and have reduced need for cooling and reduced risk of contamination). The variants may also be used in conjunction with other enzymes, particularly glucoamylase during dextrinisation. The variants are used to produce

CC syrups, particularly of high maltose content, or alcohol, from starch; as
CC dough improver for baked goods; in brewing, to increase fermentability of
CC the wort; and for liquefaction of starch

SQ Sequence 498 AA;

Query Match 100.0%; Score 2684; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 3e-226;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAMWSFLYGLQVAAPALAAPADWRSQSIYFLLTDRFARTDGGTTATCMTADQKCYCG 60
DB 1 MVAMWSFLYGLQVAAPALAAPADWRSQSIYFLLTDRFARTDGGTTATCMTADQKCYCG 60
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAGDAVHGVMQDIYSLNENYGTADDL 120
DB 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAGDAVHGVMQDIYSLNENYGTADDL 120
QY 121 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTQ 180
DB 121 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTQ 180
QY 181 VEDCWLGNNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 240
DB 181 VEDCWLGNNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 240
QY 241 KAAGVYCIGEVLDGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDDLYNNINTVK 300
DB 241 KAAGVYCIGEVLDGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDDLYNNINTVK 300
QY 301 SDPCDSTLLGTFVENHNDNPRFASYTNDIALAKNVAAFIILNDGIPIIYAGQEQHYAGGND 360
DB 301 SDPCDSTLLGTFVENHNDNPRFASYTNDIALAKNVAAFIILNDGIPIIYAGQEQHYAGGND 360
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKG 420
DB 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKG 420
QY 421 DGSQIVTILSNKGAGSDSYTILSLGAGYTAGQQLTEVIGCTTIVTGVSDGNVPPVWAGGLP 480
DB 421 DGSQIVTILSNKGAGSDSYTILSLGAGYTAGQQLTEVIGCTTIVTGVSDGNVPPVWAGGLP 480
QY 481 RVLYPTEKLAGSKICSSS 498
DB 481 RVLYPTEKLAGSKICSSS 498

RESULT 2
ADT89628

ID ADT89628 standard; protein; 499 AA.

XX AC ADT89628;

XX DT 16-DEC-2004 (first entry)

XX DE Aspergillus niger neutral alpha-amylase B (amyB) protein.

XX KW Glucoamylase; glaA; amyB; alpha-amylase B (amyB) enzyme.

XX OS Aspergillus niger.

XX PN US2004191864-A1.

XX PD 30-SEP-2004.

XX PF 31-MAR-2004; 2004US-00815495.

XX PR 31-MAR-2003; 2003US-0459902P.

XX PA (NOVO) NOVOZYMES BIOTECH INC.

XX PI Connelly M, Brody H;

XX

DR WPI; 2004-708545/69.
DR N-PSDB; ADT89627.

XX PT Producing heterologous biological substance comprises culturing mutant of
PT wild-type Aspergillus niger strain in medium suitable for producing heterologous
PT heterologous biological substance and recovering heterologous biological
PT substance.

PS Example 10; SEQ ID NO 18; 58pp; English.

XX CC The present invention relates to a method of producing heterologous
CC biological substance. The method involves culturing mutant of wild-type
CC Aspergillus niger strain in medium suitable for producing heterologous
CC biological substance, where mutant strain comprises first nucleotide
CC sequence encoding heterologous biological substance and second nucleotide
CC sequences comprising modification of glucoamylase (glaA) and recovering
CC heterologous biological substance. The present sequence is the
CC Aspergillus niger neutral alpha-amylase B (amyB) protein.

XX SQ Sequence 499 AA;

Query Match 99.6%; Score 2674; DB 8; Length 499;
Best Local Similarity 99.6%; Pred. No. 2.3e-225;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVAMWSFLYGLQVAAPALAAPADWRSQSIYFLLTDRFARTDGGTTATCMTADQKCYCG 60
DB 2 MVAMWSFLYGLQVAAPALAAPADWRSQSIYFLLTDRFARTDGGTTATCMTADQKCYCG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAGDAVHGVMQDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAGDAVHGVMQDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTQ 180
DB 122 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTQ 181
QY 181 VEDCWLGNNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCWLGNNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 241
QY 241 KAAGVYCIGEVLDGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDDLYNNINTVK 300
DB 242 KAAGVYCIGEVLDGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDDLYNNINTVK 301
QY 301 SDPCDSTLLGTFVENHNDNPRFASYTNDIALAKNVAAFIILNDGIPIIYAGQEQHYAGGND 360
DB 302 SDPCDSTLLGTFVENHNDNPRFASYTNDIALAKNVAAFIILNDGIPIIYAGQEQHYAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKG 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKG 421
QY 421 DGSQIVTILSNKGAGSDSYTILSLGAGYTAGQQLTEVIGCTTIVTGVSDGNVPPVWAGGLP 480
DB 422 DGSQIVTILSNKGAGSDSYTILSLGAGYTAGQQLTEVIGCTTIVTGVSDGNVPPVWAGGLP 481
QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499

RESULT 3
ADT89632

ID ADT89632 standard; protein; 498 AA.

XX AC ADT89632;

XX DT 16-DEC-2004 (first entry)

XX DE Aspergillus niger neutral alpha-amylase A (amyA) protein.

XX KW Glucoamylase; glaA; amyA; alpha-amylase A; enzyme.

XX Aspergillus niger.
OS US2004191864-A1.
XX 30-SEP-2004.
XX 31-MAR-2004; 2004US-00815495.
XX 31-MAR-2003; 2003US-0459902P.
XX (NOVO) NOVOZYMES BIOTECH INC.
XX Connelly M, Brody H;
XX WPI; 2004-708545/69.
DR N-PSDB; ADT89631.
XX Producing heterologous biological substance comprises culturing mutant of
PT wild-type Aspergillus niger strain in medium suitable for producing
PT heterologous biological substance and recovering heterologous biological
PT substance.
XX Example 11; SEQ ID NO 22; 58pp; English.
XX The present invention relates to a method of producing heterologous
CC biological substance. The method involves culturing mutant of wild-type
CC Aspergillus niger strain in medium suitable for producing heterologous
CC biological substance, where mutant strain comprises first nucleotide
CC sequence encoding heterologous biological substance and second nucleotide
CC sequences comprising modification of glucoamylase (glA) and recovering
CC heterologous biological substance. The present sequence is the
CC Aspergillus niger neutral alpha-amylase A (amyA) protein.
XX
SQ Sequence 498 AA;

Query Match 99.2%; Score 2662; DB 8; Length 498;
Best Local Similarity 99.8%; Pred. No. 2.6e-224;
Matches 493; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 MVAWNSLFYGLQVAAPALAAATPADWRSOSIYFLLTDRFARTDGSITATCNTADQKYCGG 60
DB 2 MVAWNSLFYGLQVAAPALAAATPADWRSOSIYFLLTDRFARTDGSITATCNTADQKYCGG 61

QY 61 TWGGIIDLKDYIQGMGFTAIWITPVTQAQLPQTAYGDYHGYWQDIYSLNENYGTADDL 120
DB 62 TWGGIIDLKDYIQGMGFTAIWITPVTQAQLPQTAYGDYHGYWQDIYSLNENYGTADDL 121

QY 121 KALSSALHERGMYLMDVNVANHMGYDGAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQ 180
DB 122 KALSSALHERGMYLMDVNVANHMGYDGAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQ 181

QY 181 VEDCWLGDNTVSLPDLDTTKDVVKNEWYDMVGLSVNSYIDGLRIDTVKHVQKDFWPGYN 240
DB 182 VEDCWLGDNTVSLPDLDTTKDVVKNEWYDMVGLSVNSYIDGLRIDTVKHVQKDFWPGYN 241

QY 241 KAAGVYICIGELVDGDPAYTCYQNVMDGVNLNPIYYPYLLNAPKSTSGSMDLLNMLNTVK 300
DB 242 KAAGVYICIGELVDGDPAYTCYQNVMDGVNLNPIYYPYLLNAPKSTSGSMDLLNMLNTVK 301

QY 301 SDCPDSTLLGTFFVENHNDNPRFASYTNDIALAKNVAAPFIILNDGIPPIYAGQEQHYAGND 360
DB 302 SDCPDSTLLGTFFVENHNDNPRFASYTNDIALAKNVAAPFIILNDGIPPIYAGQEQHYAGND 361

QY 361 PANREATWLSGYPTDSSELYKLIASANAIRNYAISKDTGFVYKQWPIYKDDTTIAMRKG 420
DB 362 PANREATWLSGYPTDSSELYKLIASANAIRNYAISKDTGFVYKQWPIYKDDTTIAMRKG 421

QY 421 DGSQIVTILSNKGASGDSYTLISLGAGYTAGQQLTEVIGCTTVTVGSDGNVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGAGYTAGQQLTEVIGCTTVTVGSDGNVPVPMAGGLP 481

QY 481 RVLYPEKLAGSKIC 495

DB 482 RVLYPEKLAGSKIC 496

RESULT 4
AAR72450
ID AAR72450 standard; protein; 478 AA.
XX AAR72450;
AC AAR72450;
XX 25-MAR-2003 (revised)
DT 01-DEC-1995 (first entry)
XX
DE Aspergillus oryzae alpha amylase (mature protein).
XX Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
KW thermostable.
XX Aspergillus oryzae.
OS W09510603-A1.
XX
PN 20-APR-1995.
PD 05-OCT-1994; 94WO-DK000370.
XX 08-OCT-1993; 93DK-00001133.
PR 02-FEB-1994; 94DK-00000140.
XX (NOVO) NOVO-NORDISK AS.
XX
PI Borchert TV, Bisgard-Frantzen H, Svendsen A, Thellersen M;
PI Van Der Zee P;
XX WPI; 1995-161790/21.
DR New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance.
PT Disclosure; Page 75-76; 105pp; English.
PS Variant alpha amylase enzymes which have improved washing and/or as
XX detergent additives. The enzymes have one or more amino acid residues
CC added, deleted or substituted. The variants can also be used for textile
CC desizing prior to scouring, bleaching and dyeing. The variants have
CC improved thermostability, acid/alkaline stability; low temperature
CC optimum; pH optimum; higher hydrolysis velocity and improved tolerance to
CC other composition constituents, e.g. oxidation agents. (Updated on 25-MAR
CC -2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 478 AA;

Query Match 95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.9e-216;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSOSIYFLATDRPARTDGSITATCNTADQKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSOSIYFLATDRPARTDGSITATCNTADQKYCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTQAQLPQTAYGDYHGYWQDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 140
DB 61 WITPVTQAQLPQTAYGDYHGYWQDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 120

QY 141 NHMGYDGAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTVQVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDGAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTVQVEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDMVGLSVNSYIDGLRIDTVKHVQKDFWPGYNKAAGVYICIGELVDGDPAYTC 260
DB 181 DVVKNEWYDMVGLSVNSYIDGLRIDTVKHVQKDFWPGYNKAAGVYICIGELVDGDPAYTC 240

```
QY 261 PYQNVMDGVLNPIYPIYPLNNAFKSTSGSMDDLYNNINTVKSQPCDSTLLGTFVFNHNDPR 320
DB 241 PYQNVMDGVLNPIYPIYPLNNAFKSTSGSMDDLYNNINTVKSQPCDSTLLGTFVFNHNDPR 300
QY 321 FASYTNDIALAKNVAAPFIILNDGIPIIYAGQHQHVAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAPFIILNDGIPIIYAGQHQHVAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNYAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNYAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420
QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 5
AAR78270
ID AAR78270 standard; protein; 478 AA.
XX
AC AAR78270;
XX
DT 17-JAN-1996 (first entry)
XX
DE Aspergillus oryzae alpha amylase (mature protein).
XX
KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric; starch;
KW thermostable; methionine; Bacillus licheniformis;
KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;
KW Aspergillus oryzae.
XX
OS Aspergillus oryzae.
XX
PN WO9521247-A1.
XX
PD 10-AUG-1995.
XX
PF 05-OCT-1994; 94WO-DK000371.
XX
PR 02-FEB-1994; 94DK-00000141.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Toft AH, Marcher D, Pedersen HH, Nilsson TE;
XX
WPI; 1995-283767/37.
XX
DR Use of an oxidation stable alpha-amylase - for simultaneous desizing and
PT bleaching or scouring of fabrics contg. starch or starch derivs.
XX
PS Disclosure; Page 25-26; 37pp; English.
XX
CC Oxidation stable alpha amylases can be used for the simultaneous desizing
CC and bleaching or scouring of a fabric comprising starch or starch
CC derivatives. They exhibit a better heat stability, especially in the
CC presence of oxidizing agents. They are obtained from a parent alpha
CC amylase by replacing one or more methionine residues with any amino acid
CC different from Cys or Met, preferably Leu, Thr, Ala, Gly, Ser, Ile or
CC Asp. The parent alpha amylase is pref. derived from a Bacillus species,
CC although alpha amylases of fungal origin can also be used. This sequence
CC is the wild type (unmodified) alpha amylase of Aspergillus oryzae
XX
SQ Sequence 478 AA;

Query Match 95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.9e-216;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLTLTDRPARTDGTSTATCNDQKYCGGTWQGIIDKLDYIQSGMFTAI 80
DB 1 ATPADWRSQSIYFLTLTDRPARTDGTSTATCNDQKYCGGTWQGIIDKLDYIQSGMFTAI 60
```

```
QY 81 WITPVTALPQTATGDAYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTALPQTATGDAYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120
QY 141 NHMGYDAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTOVEDCWLGNTVSLPDLDTTK 200
DB 121 NHMGYDAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTOVEDCWLGNTVSLPDLDTTK 180
QY 201 DVVKNWYDWVGSLSVSNYSIDGLRIDTVKHQKDFWPGYNKAAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNWYDWVGSLSVSNYSIDGLRIDTVKHQKDFWPGYNKAAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVLNPIYPIYPLNNAFKSTSGSMDDLYNNINTVKSQPCDSTLLGTFVFNHNDPR 320
DB 241 PYQNVMDGVLNPIYPIYPLNNAFKSTSGSMDDLYNNINTVKSQPCDSTLLGTFVFNHNDPR 300
QY 321 FASYTNDIALAKNVAAPFIILNDGIPIIYAGQHQHVAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAPFIILNDGIPIIYAGQHQHVAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNYAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNYAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420
QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 6
AAW14500
ID AAW14500 standard; protein; 478 AA.
XX
AC AAW14500;
XX
DT 04-JUN-1997 (first entry)
XX
DE Aspergillus oryzae alpha-amylase (mature protein).
XX
KW alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;
KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
KW calcium dependency; substrate binding; stability; pH optimum;
KW thermostability; cleavage; oligosaccharide substrate; dishwashing;
KW washing; detergent additive; fabric desizing; starch liquefaction;
KW sweetener; ethanol production; variant.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Misc-difference 13..45
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent alpha
FT -amylase (used as a template for a variant) corresponding
FT to 7-23 of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim 33"
FT Misc-difference 14..40
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a template for
FT a variant) corresponding to 8-18 of AAW14499 is deleted
FT or replaced with a fragment corresponding to this
FT fragment; claim 35"
FT Misc-difference 28..42
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent alpha
FT -amylase (used as a template for a variant) corresponding
FT to 12-19 of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim 30"
FT Misc-difference 32..38
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a template for
```

FT	a variant) corresponding to 14-15 of AAW14499 is deleted
FT	or replaced with a fragment corresponding to this
FT	fragment; claim 32"
FT	66. .84
FT	/label= loop 2 modification region
FT	/note= "at least one amino acid residue of a parent alpha
FT	-amylase (used as a template for a variant) corresponding
FT	to 44-57 of AAW14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim 18"
FT	70. .78
FT	/label= loop 2 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a template for
FT	a variant) corresponding to 48-51 of AAW14499 is deleted
FT	or replaced with a fragment corresponding to this
FT	fragment; claim 20"
FT	98. .210
FT	/label= loop 3 modification region
FT	/note= "at least one amino acid residue of a parent alpha
FT	-amylase (used as a template for a variant) corresponding
FT	to 117-185 of AAW14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim 24"
FT	102. .206
FT	/label= loop 3 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a template for
FT	a variant) corresponding to 121-181 of AAW14499 is
FT	deleted or replaced with a fragment corresponding to this
FT	fragment; claim 26"
FT	121. .181
FT	/note= "at least one amino acid residue of a parent alpha
FT	-amylase (used as a template for a variant) corresponding
FT	to this fragment is deleted or replaced with a fragment
FT	corresponding to 102-206 of AAW14499; claim 41"
FT	121. .174
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a template for
FT	a variant) corresponding to this fragment is deleted or
FT	replaced with a fragment corresponding to 102-199 of
FT	AAW14499; claim 42"
FT	165. .177
FT	/label= loop 3 modification region
FT	/note= "at least one amino acid residue of a parent alpha
FT	-amylase (used as a template for a variant) corresponding
FT	to 195-202 of AAW14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim 21"
FT	166. .173
FT	/label= loop 3 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a template for
FT	a variant) corresponding to 196-198 of AAW14499 is
FT	deleted or replaced with a fragment corresponding to this
FT	fragment; claim 23"
FT	181. .184
FT	/note= "an amino acid fragment corresponding to this
FT	region is deleted from the parent sequence of a variant
FT	Funcamyl; claim 43"
FT	291. .313
FT	/label= loop 8 modification region
FT	/note= "at least one amino acid residue of a parent alpha
FT	-amylase (used as a template for a variant) corresponding
FT	to 322-346 of AAW14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim 36"
FT	297. .313
FT	/label= loop 8 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a template for
FT	a variant) corresponding to 325-345 of AAW14499 is
FT	deleted or replaced with a fragment corresponding to this
FT	fragment; claim 38"
XX	
XX	WO9623874-A1.
XX	
XX	

AC ABP96630;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Alpha-amylase/glucoamylase fusion protein sequence SEQ ID NO:45.
 XX
 KW self-processing plant; plant; processing enzyme; alpha-amylase; grain;
 KW pullulanase; alpha-glucosidase; glucose isomerase; glucoamylase;
 KW mesophilic; thermophilic; hyperthermophilic; transgenic plant; starch;
 KW maltodextrin; ethanol; fermentation; beverage; enzyme.
 XX
 OS Aspergillus shirousami.
 OS Synthetic.
 PN WO2003018766-A2.
 XX
 PD 06-MAR-2003.
 XX
 PF 27-AUG-2002; 2002WO-US027129.
 XX
 PR 27-AUG-2001; 2001US-0315281P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Lanahan MB, Basu SS, Batie CJ, Chen W, Craig J, Kinkema M;
 DR WPI; 2003-268420/26.
 DR N-PSDB; ACC44572.
 XX
 XX Novel polynucleotide encoding hyperthermophilic processing enzymes e.g.
 PT alpha-amylase, useful for producing plant to produce food products having
 PT improved taste or fermentable substrates for ethanol.
 XX
 PS Claim 1; Page 107; 158pp; English.
 XX
 CC The present invention describes polynucleotides which encode processing
 CC enzymes (e.g. alpha-amylase, pullulanase, alpha-glucosidase, glucose
 CC isomerase, or glucoamylase) that are optimised for expression in plants.
 CC The polynucleotides encode mesophilic, thermophilic or hyperthermophilic
 CC processing enzymes, which are activated under suitable conditions to act
 CC upon the desired substrate. Also described are self-processing transgenic
 CC plants and plant parts, e.g. grain, which express one or more of these
 CC enzymes and have an altered composition that facilitates starch and grain
 CC processing. Also described is a method (M) for converting starch to
 CC starch-derived products in a transformed plant part (tpp), by activating
 CC the starch processing enzyme contained in it. Transgenic grain is useful
 CC for preparing maltodextrin. A transformed plant (TP) can be used to
 CC produce food products having improved taste and to produce fermentable
 CC substrates for ethanol and fermented beverages. (M) eliminates the need
 CC to mill or physically disrupt the integrity of plant parts prior to
 CC recovery of starch-derived products. The present sequence represents
 CC alpha-amylase/glucoamylase fusion protein, which is given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 1095 AA;
 Query Match 95.7%; Score 2569; DB 6; Length 1095;
 Best Local Similarity 99.8%; Pred. No. 1.3e-215;
 Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 21 ATPADWRSQSIYFLLDREPARTDGGTTATCNCADQKCGGTWQGIIDKLDYIQSGGFTAI 80
 DB 1 ATPADWRSQSIYFLLDREPARTDGGTTATCNCADQKCGGTWQGIIDKLDYIQSGGFTAI 60
 QY 81 WITPTAQLPQTAYGDAVHGYYQQDIYSLNENYGTADDLKALSSALHGRMYLMDVVA 140
 DB 61 WITPTAQLPQTAYGDAVHGYYQQDIYSLNENYGTADDLKALSSALHGRMYLMDVVA 120
 QY 141 NHMGVDGAGSSVDYVFKPSSQDYFHPFCFQNYEDQTVQEDCWLGDNTVSLPDLDTTK 200
 DB 121 NHMGVDGAGSSVDYVFKPSSQDYFHPFCFQNYEDQTVQEDCWLGDNTVSLPDLDTTK 180
 QY 201 DVVKNWYDWVGLSVNSYSDGLRIDTVKRVQKDFWPGYNKAAGVYICIGVLDGDPAYTC 260

DB 181 DVVKNWYDWVGLSVNSYSDGLRIDTVKRVQKDFWPGYNKAAGVYICIGVLDGDPAYTC 240
 QY 261 PYQNVMDGLNYPYIYPLLNAFKSTGSMDDLYNMINTVKSDCPDSTLLGTFEVHNDR 320
 DB 241 PYQNVMDGLNYPYIYPLLNAFKSTGSMDDLYNMINTVKSDCPDSTLLGTFEVHNDR 300
 QY 321 FASYTNDIALAKNVAAFIILNDGPIIYAGQEHYAGNDPANREATWLSGYPTDSELYK 380
 DB 301 FASYTNDIALAKNVAAFIILNDGPIIYAGQEHYAGNDPANREATWLSGYPTDSELYK 360
 QY 381 LIASANAIRNVAISKDTGFVTKWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSVT 440
 DB 361 LIASANAIRNVAISKDTGFVTKWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSVT 420
 QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
 DB 421 LSLSGAGYTAGQQLTEVIGCTTIVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478
 RESULT 8
 AAR46065
 ID AAR46065 standard; protein; 478 AA.
 XX
 AC AAR46065;
 XX
 DT 25-MAR-2003 (revised)
 DT 18-JUL-1994 (first entry)
 XX
 DE Mutant alpha-amylase.
 XX
 KW Methionine substitution; stability; activity; detergent;
 KW dishwashing agents; liquifaction agents.
 XX
 OS Aspergillus oryzae.
 XX
 PN WO9402597-A1.
 XX
 PD 03-FEB-1994.
 XX
 PF 06-JUL-1993; 93WO-DK000230.
 XX
 PR 23-JUL-1992; 92DK-00000946.
 PR 16-DEC-1992; 92DK-00001503.
 PR 15-MAR-1993; 93DK-00000292.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Svendsen A, Bisgard-Frantzen H;
 XX
 DR WPI; 1994-048855/06.
 XX
 PT Mutant alpha-amylase from Bacillus species comprising a methionine
 PT substitution - with improved stability and activity at low pH, for use in
 PT detergents, dishwashing agents and liquifaction agents.
 XX
 PS Claim 1; Page 7; 20pp; English.
 XX
 CC The sequence os that of the Aspergillus oryzae alpha amylase, sold
 CC commercially as FUNGAWYL (TM) by Novo Nordisk A/S. The sequence can be
 CC mutated by substitution of one or more of its methionine residues for any
 CC amino acid other than cysteine. The mutant alpha-amylase exhibits a
 CC better activity level and better stability in the presence of oxidising
 CC agents than previous mutant alpha amylases, and improved thermostability
 CC at moderately low pH. The enzyme can be used as an additive for
 CC detergents, dishwashing agents and liquifaction agents. (Updated on 25-
 CC MAR-2003 to correct PN field.)
 XX
 SQ Sequence 478 AA;
 Query Match 95.1%; Score 2552; DB 2; Length 478;
 Best Local Similarity 99.4%; Pred. No. 1.1e-214;
 Matches 475; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYELLTDREARTDGTATCNTADQKFCGGTGWGIIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYELLTDREARTDGTATCNTADQKFCGGTGWGIIIDKLDYIQGMGFTAI 60
QY 81 WITPVTQAQLPQTTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTQAQLPQTTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120
QY 141 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVTVSLPDLDTTK 200
DB 121 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVTVSLPDLDTTK 180
QY 201 DVVKNEWYDWVGLSVNSYSDGLRIDTVKHVQKDFWPGYNKAAGVYICIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDWVGLSVNSYSDGLRIDTVKHVQKDFWPGYNKAAGVYICIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVLNPIYYPYLLNAPKSTSGMDDLNNMINTVKSDCPDSTLLGTTFVENHDNPR 320
DB 241 PYQNVMDGVLNPIYYPYLLNAPKSTSGMDDLNNMINTVKSDCPDSTLLGTTFVENHDNPR 300
QY 321 FASYTNDIALAKNVAAFIILNDGPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420
QY 441 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPPVMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPPVMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 9

ID AAR79025 standard; protein; 478 AA.

AC AAR79025;

DT 22-MAR-1996 (first entry)

DE Mature taka-amylose A.

XX Wild type; neopullulanase; B. stearotheophilus; mutant; food industry; modification; hydrophobicity; replacement insertion; deletion.

XX Aspergillus oryzae.

XX Key Location/Qualifiers

FT Disulfide-bond 30..38

FT Disulfide-bond 150..164

FT Disulfide-bond 240..283

FT Disulfide-bond 439..474

XX JP07177891-A.

XX 18-JUL-1995.

XX 31-OCT-1994; 94JP-00288658.

XX 12-NOV-1993; 93JP-00306096.

XX (NIDE) NEC CORP.

XX (EZAK) EZAKI GLICO CO.

XX WPI; 1995-279919/37.

XX Modifying a transferase by enhancing hydrophobicity of a selected site - increases transfer activity, also new mutant neo-pullulanase(s).

PS Disclosure; Page 10-11; 18pp; English.

XX

CC This sequence represents the mature form of taka-amylose from A. oryzae.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the invention
CC for enhancing the hydrophobicity of a selected site of the pullulanase.
CC The method comprises replacement of a group in the selected site with a
CC hydrophobic group, replacement of an amino acid with a hydrophobic amino
CC acid, and/or insertion or deletion of a hydrophobic amino acid from the
CC S422V and M375L
XX

SQ Sequence 478 AA;

Query Match 93.3%; Score 2503.5; DB 2; Length 478;
Best Local Similarity 97.9%; Pred. No. 28-210;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 21 ATPADWRSQSIYELLTDREARTDGTATCNTADQKFCGGTGWGIIIDKLDYIQGMGFTAI 80

DB 1 ATPADWRSQSIYELLTDREARTDGTATCNTADQKFCGGTGWGIIIDKLDYIQGMGFTAI 60

QY 81 WITPVTQAQLPQTTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140

DB 61 WITPVTQAQLPQTTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120

QY 141 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVTVSLPDLDTTK 200

DB 121 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVTVSLPDLDTTK 180

QY 201 DVVKNEWYDWVGLSVNSYSDGLRIDTVKHVQKDFWPGYNKAAGVYICIGEVLDGDPAYTC 260

DB 181 DVVKNEWYDWVGLSVNSYSDGLRIDTVKHVQKDFWPGYNKAAGVYICIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYYPYLLNAPKSTSGMDDLNNMINTVKSDCPDSTLLGTTFVENHDNPR 320

DB 241 PYQNVMDGVLNPIYYPYLLNAPKSTSGMDDLNNMINTVKSDCPDSTLLGTTFVENHDNPR 300

QY 321 FASYTNDIALAKNVAAFIILNDGPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 380

DB 301 FASYTNDIALAKNVAAFIILNDGPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440

DB 361 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 419

QY 441 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPPVMAGGLPRVLYPTEKLAGSKICSSS 498

DB 420 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPPVMAGGLPRVLYPTEKLAGSKICSSS 477

RESULT 10

ABB09072

ID ABB09072 standard; protein; 423 AA.

XX ABB09072;

XX 26-JUN-2002 (first entry)

XX Aspergillus oryzae TAKA protein (TAA).

XX Thermus sp. IM6501; maltogenic amylase; EC 3.2.1.133; crystallisation;
KW protein co-ordinate data; 3 dimensional structure.

XX Aspergillus oryzae.

XX KR2001027418-A.

XX 06-APR-2001.

XX 09-SEP-1999; 99KR-00039130.

XX 09-SEP-1999; 99KR-00039130.

XX

CC patterns, and to raise anti-protein antibodies. The polypeptide having
CC tannase activity increases the yield of tea liquor from tea leaves,
CC improves the colour, flavour and health benefits of tea products,
CC particularly an instant tea product. The polypeptide having cellulase
CC activity enhances cleaning ability of detergent compositions

XX
SQ Sequence 495 AA;

	Query Match	78.2%; Score 2100; DB 6; Length 495;
	Best Local Similarity	78.6%; Pred. No. 5.3e-175;
	Matches 389; Conservative	39; Mismatches 65; Indels 2; Gaps 1;

QY 2 VAWNSFLYLGLVAAPALAAPADWRSSSIYFLLTDRPARTDGGSTATCNTADOKYCGGT 61
DB : |||||
3 IRWSFFLSCL--AGTALAATPAQRSSSIYFLLTDRPARTDGGSTTASCDSAREYCGGT 60
QY 62 WQGIIDKLDYIQGFGTAIWTPVTAQLPOTTAYGDVHGVMQDIYSLNENYGTADDLK 121
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 61
WQGIIEQDYIYQGFGTAIWTPVKLPQTSSGTAYHGMQDIYSVNSNYGTADDLK 120
QY 122 ALSSALHERGMVLMDVVANHWGVDGAGSSVDYSVFPPFSQDYFHPFCFIQNYEDOTQV 181
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121
ALASALHDRGMVLMDVVANHWGYPAGAGSDVDYSVFNPFSQTSFHPLCFISINYDNQTDV 180
QY 182 EDCWLGNNTVSLPDLTDTTKVGVKNWDMVGSLVSNSYIDGLRIDTVKHVQKDFWPYNK 241
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181
ENCWLGNVSPLPDLTDTNPDPVKIWNWNSLVSNYSIDGLRIDTVKHVQSDFWPGFND 240
QY 242 AAGVYCIGEVLDDGPATCTPCYQNTWDGVLANPYIYYPLLNAFKSTSGSMDDLNNINTVKS 301
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241
AAGVYCIGEVDGDPATCTCPQEVLDGVLNPIYYPLLKAFQSTSGMSLSLYMDINTVKS 300
QY 302 DCPDSTLLTGTFVENHDNPRFASYTNIDIALAKVAAFTILNDGIPIIYAGQBHQHVAGGNDP 361
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301
QCADSTLLTGTFVENHDTFRFASYTKDWALAKVAAAFIFSDCIPIIYAGQBHQHYSGGADP 360
QY 362 ANREATWLSGYPTDSLEYKLIASANAIRNVAISKDGTGFVTYKKNWPIYKDDTTIAMRRKTD 421
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 361
ANREAVWLSGYSTSDLYKLIATANAIRSHAISKDPGVITYKKNPIYKDTSTIAMRKGS 420
QY 422 GSQIVTILSKAGSDSYTLISLGSAGYTACQOLTVEIGCTTITVVGSDGNVPVPNAGELPR 481
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 421
GAQIITVLSNLGASGSSYTLISLGTYEAGQQLTMEFSCCTTITVVGSDKKVPVMSAGLPR 480
QY 482 VLYPTEKLAGSKICS 496
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 481
VFYPTAGLNGSTVCT 495

RESULT 12
ABP97894
ID ABP97894 standard; protein; 494 AA.
XX AC
XX ABP97894;
XX DT
XX 17-JUN-2003 (first entry)
XX DE
XX Amino acid sequence of an alpha-amylase of Aspergillus niger.
XX DE
XX Alpha-amylase; enzyme; endohydrolysis; 1,4-alpha-glucosidic linkage;
KW oligosaccharide; polysaccharide; baking.
XX KW
XX Aspergillus niger.
OS
XX
XX WO2003016535-A2.
PN
XX
XX 27-FEB-2003.
XX
XX
XX 02-AUG-2002; 2002WO-NL000522.
PF
XX
XX 16-AUG-2001; 2001EP-00000379.
PR 16-AUG-2001; 2001EP-00000380.
PR 16-AUG-2001; 2001EP-00000381.

Enzyme; tannase; cellulase; glucose oxidase; glucoamylase; phytase; beta-galactosidase; invertase; lipase; alpha-amylase; laccase; polygalacturonase; xylanase; gallate ester linkage; detergent; cellulose; glucose; oxygen; myo-inositol phosphate; laccase; tea leaf; sucrose; glyceride; starch; maltodextrin; oxidized phenolic compound; polygalacturonic acid chain; xylan; xylo-oligomer; food; feed; beverage; textile; tea liquor; cleaning ability.

Aspergillus fumigatus.

WO2003012071-A2.

13-FEB-2003.

05-AUG-2002; 2002WO-US024842.

03-AUG-2001; 2001US-0309870P.

(ELIT-) ELITRA PHARM INC.

Jiang B, Storms R, Roemer T, Bussey H; WPI; 2003-332729/31.

N-PSDB; ABQ80347, ABQ80348.

Novel isolated Aspergillus fumigatus polypeptide, useful in various industries such as those involved in the making of food and feed, beverages, textiles and detergents.

Claim 17; Page 139-40; 169pp; English.

The sequences given in AB80164-87 show enzymatic proteins derived from A. fumigatus. These proteins display the catalytic activity of an enzyme such as tannase, cellulase, glucose oxidase, glucoamylase, phytase, beta-galactosidase, invertase, lipase, alpha-amylase, laccase, polygalacturonase or xylanase. Compositions comprising the tannase are useful for modulating the amount of compounds that comprise a gallate ester linkage in a composition. Compositions comprising cellulase are useful for modulating the amount of cellulose in a composition. Compositions comprising glucose oxidase are useful for modulating the amount of glucose or oxygen in a composition. Compositions comprising phytase are useful for modulating the amount of myo-inositol phosphates in a composition. Compositions comprising beta-galactosidase are useful for modulating the amount of lactose in a composition. Compositions comprising sucrose or invertase are useful for modulating the amount of sucrose in a composition. Compositions comprising lipase are useful for modulating the amount of glyceride in a composition. Compositions comprising alpha-amylases are useful for modulating the amount of starches or maltodextrins in a composition. Compositions comprising laccase are useful for modulating the amount of oxidized phenolic compounds in a composition. Compositions comprising polygalacturonases are useful for modulating the amount of high or low molecular weight polygalacturonic acid chains in a composition. Compositions comprising xylanases are useful for modulating the amount of xylan or xylo-oligomers in a composition. The A. fumigatus proteins and corresponding DNA's are useful in various industries such as those involved in the making of food and feed, beverages, textiles and detergents. The DNA's are useful to express recombinant enzymes for characterization, modification or industrial uses, to compare with the nucleotide sequence of A.fumigatus to identify duplicated genes of paralogs having the same or similar biochemical activity and/or function, to compare with nucleic acid sequence of other related or distant fungal organisms to identify potential orthologous enzyme genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, and to raise anti-protein antibodies. The polypeptide having tannase activity increases the yield of tea liquor from tea leaves, improves the colour, flavour and health benefits of tea products, particularly an instant tea product. The polypeptide having cellulase activity enhances cleaning ability of detergent compositions

Sequence 630 AA;

Query Match 68.4%; Score 1835; DB 6; Length 630;
Best Local Similarity 66.8%; Pred. No. 1.4e-151;
Matches 334; Conservative 59; Mismatches 101; Indels 6; Gaps 2;

QY 5 WSLFLYGLQVAAPAL-----RATPADWRSQSIFLLTDRPARTDGTTCATNTAQQKYG 59
DB 3 WISQLFPLSLCSLQGAHALTPAERWSQSIFLLTDRFREDNSTTAACDVTQRLYCG 62
QY 60 GTWQGIIDKLDYIQMGFTAIWITVTAQLPQTAYGDAYHGYWQDYSLNENYGTAD 119
DB 63 GSWQGIINHLDYIQMGFTAIWITVTEQFGDTSYHGYWQONHEVNANVGTAD 122
QY 120 LKALSSALHERGMYLWVDVANHMGYDAGAGSDVSVFKPSSQDYFHPFCFIQNYEDQT 179
DB 123 LRDLANALHARGMYLWVDVANHMGYNGAGSNVYGVTFPPDSATYFHPYCLITDYNQT 182
QY 180 QVEDCWLGDNTVSLPDLDTTKDVKVNEWYDVGSLVSNYSIDGLRIDTVKHVKQDFWPGY 239
DB 183 AVEDCWLGDNTVSLPDLDTTKDVKVNEWYDVGSLVSNYSIDGLRIDTVKHVKQDFWPGY 242
QY 240 NKAAGVYCI GEVL DGDPRATCPYQNVMDGVNLNPIYYPILLNAFKSTSGSMDLYNNINTV 299
DB 243 NDAAGVYCVGEVFGDPQYTCYQNYLDGVNLNPIYYPILLNAFKSTSGSMDLYNNINTV 302
QY 300 KSDCPDSTLLGTFFVENHDNPRFASVTNDIALAKNVAFFIILNDGPIIYVAGQEHYAGN 359
DB 303 ASDCADPTLLGNFIENHDNPRFASVTSDVQAKNWFPMFSDGPIVYVAGQEHYAGN 362
QY 360 DPANREATWLSGYPDSELYKLIASANAIRNVAISKDTGFTVYKWPYKDDTTAMRKG 419
DB 363 DPANREAVWLSGYSTSATLYSWIASTNKRKLAISKDSAYITSKNPFYDSTNTLAMRKG 422
QY 420 T-DGSOIVTILSNKAGSDSYTSLSGAGYTAGQQLTEVIGCTTVTGSDGNVPVPMAGG 478
DB 423 SVAGSQIVTILSNKAGSDSYTSLSGAGYTAGQQLTEVIGCTTVTGSDGNVPVPMAGG 482
QY 479 LPRVLYPTEKLAGSKICSSS 498
DB 483 LPRVFPVSSWVSGSLCGDS 502

RESULT 14
AAE24207
ID AAE24207 standard; protein; 484 AA.
AC AAE24207;
XX
DT 04-OCT-2002 (first entry)
XX
DE Aspergillus niger alpha-amylase protein.
XX
KW Ethanol production; starch; fermentation; liquefaction; alpha-amylase;
KW fuel alcohol; fuel additive; neutral spirit; industrial ethanol; enzyme.
XX
OS Aspergillus niger.
XX
PN WO200238787-A2.
XX
PD 16-MAY-2002.
XX
PF 09-NOV-2001; 2001WO-DK000737.
XX
PR 10-NOV-2000; 2000DK-00001676.
PR 21-NOV-2000; 2000US-0252213P.
PR 11-DEC-2000; 2000DK-00001854.
PR 15-DEC-2000; 2000US-0256015P.
XX
PA (NOVO) NOVOZYMES AS
PA (NOVO) NOVOZYMES NORTH AMERICA INC.
XX
PI Veit C, Felby C, Fuglsang CC;
XX
DR WPI; 2002-479793/51.

Search completed: February 8, 2005, 18:36:39
Job time : 293.866 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:29:19 ; Search time 76.9112 Seconds
(without alignments)
483.353 Million cell updates/sec

Title: US-10-820-200-2

Perfect score: 2684

Sequence: 1 MWANSLFLYGLQVAPALA.....LPRVLPTEKLAGSKICSS 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2572	95.8	478	1	US-08-720-899-7
2	2572	95.8	478	1	US-08-459-610-7
3	2572	95.8	478	2	US-08-343-804-7
4	2572	95.8	478	2	US-08-600-908A-10
5	2572	95.8	478	3	US-08-683-838A-10
6	2572	95.8	478	3	US-09-182-859-7
7	2572	95.8	478	4	US-09-572-459-7
8	2572	95.8	478	4	US-09-536-252A-10
9	2572	95.8	478	4	US-10-186-042-7
10	2503.5	93.3	478	2	US-08-339-715A-2
11	1392	51.9	468	1	US-08-470-702-6
12	1392	51.9	468	1	US-08-467-831-6
13	1391	51.8	468	1	US-08-204-656B-4
14	1391	51.8	468	1	US-08-470-702-7
15	1391	51.8	468	1	US-08-467-831-7
16	1388	51.7	468	1	US-08-204-656B-6
17	1388	51.7	468	1	US-08-470-702-8
18	1388	51.7	468	1	US-08-467-831-8
19	1387	51.7	468	1	US-08-204-656B-2
20	1387	51.7	468	1	US-08-204-656B-8
21	1387	51.7	468	1	US-08-470-702-9
22	1387	51.7	468	1	US-08-467-831-9
23	450	16.8	719	3	US-09-386-607-2
24	450	16.8	719	4	US-09-645-707B-2
25	446	16.6	686	3	US-08-947-965-73
26	444	16.5	685	3	US-08-947-965-72
27	428.5	16.0	655	1	US-08-469-202-28

28 428.5 16.0 655 2 US-08-484-434C-35 Sequence 35, Appl
29 428.5 16.0 655 4 US-09-384-361-37 Sequence 35, Appl
30 427.5 15.9 655 1 US-08-469-202-27 Sequence 37, Appl
31 427.5 15.9 655 2 US-08-484-434C-34 Sequence 34, Appl
32 427.5 15.9 655 4 US-08-484-434C-34 Sequence 34, Appl
33 426.5 15.9 683 3 US-08-947-965-2 Sequence 2, Appl
34 426.5 15.9 683 3 US-08-947-965-2 Sequence 2, Appl
35 421.5 15.7 624 3 US-08-947-965-78 Sequence 78, Appl
36 418 15.6 676 3 US-08-947-965-71 Sequence 71, Appl
37 413 15.4 686 3 US-08-947-965-70 Sequence 70, Appl
38 404.5 15.1 833 3 US-08-947-965-75 Sequence 75, Appl
39 404.5 15.1 833 4 US-10-014-436-3 Sequence 3, Appl
40 404.5 15.1 1938 3 US-09-514-302-2 Sequence 2, Appl
41 404.5 15.1 1938 4 US-10-014-436-2 Sequence 2, Appl
42 398 14.8 685 3 US-08-947-965-74 Sequence 74, Appl
43 396 14.8 675 3 US-08-947-965-76 Sequence 76, Appl
44 385.5 14.4 680 3 US-08-947-965-77 Sequence 77, Appl
45 370 13.8 588 2 US-08-339-715A-5 Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 95.8%; Score 2572; DB 1; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRPARTGSGTATCANTADQKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRPARTGSGTATCANTADQKYCGGTWQGIIDKLDYIQGMGFTAI 60
QY 81 WITPVTAAQLPQTAYGDAYHGWWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTAAQLPQTAYGDAYHGWWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120
QY 141 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGDNTVSLPDLDTTK 180
QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVNLPIIYYPILLNAFKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHDNPR 320
DB 241 PYQNVMDGVNLPIIYYPILLNAFKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHDNPR 300
QY 321 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNYAISKDTGFVTKNWP IYKDDTTIAMRKGTGDSQIVTILSNKAGSGDSYT 440
DB 361 LIASANAIRNYAISKDTGFVTKNWP IYKDDTTIAMRKGTGDSQIVTILSNKAGSGDSYT 420
QY 441 LSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 2

US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043o No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-610-7
Query Match 95.8%; Score 2572; DB 1; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 1; Mismatches 0; Gaps 0;
QY 21 ATPADWRSQSIYFLLTDRPARTGSGTATCANTADQKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRPARTGSGTATCANTADQKYCGGTWQGIIDKLDYIQGMGFTAI 60
QY 81 WITPVTAAQLPQTAYGDAYHGWWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTAAQLPQTAYGDAYHGWWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120
QY 141 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGDNTVSLPDLDTTK 180
QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVNLPIIYYPILLNAFKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHDNPR 320
DB 241 PYQNVMDGVNLPIIYYPILLNAFKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHDNPR 300
QY 321 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNYAISKDTGFVTKNWP IYKDDTTIAMRKGTGDSQIVTILSNKAGSGDSYT 440
DB 361 LIASANAIRNYAISKDTGFVTKNWP IYKDDTTIAMRKGTGDSQIVTILSNKAGSGDSYT 420
QY 441 LSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478
RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5830837o No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.

```
;
;
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054,214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: linear
;
US-08-343-804-7

Query Match          95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGTATCTATCNDOKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRFARTDGTATCTATCNDOKYCGGTWQGIIDKLDYIQGMGFTAI 60
QY 81 WITPVTALPQTAYGDAYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTALPQTAYGDAYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120
QY 141 NHMGVDGAGSSVDYVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNVSLPDLDTTK 200
DB 121 NHMGVDGAGSSVDYVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNVSLPDLDTTK 180
QY 201 DVVKNEWYDMWGLSVNSYIDGLRIDTVKHVKQDFMPGYNKAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDMWGLSVNSYIDGLRIDTVKHVKQDFMPGYNKAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSSTLLGTTFVENHNDPR 320
DB 241 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSSTLLGTTFVENHNDPR 300
QY 321 FASYTNDIALAKNVAAFIILNDGIPPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420
QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 4
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5989169o No. 5989169disk of No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
;
;
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-600-908A-10

Query Match          95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGTATCTATCNDOKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRFARTDGTATCTATCNDOKYCGGTWQGIIDKLDYIQGMGFTAI 60
QY 81 WITPVTALPQTAYGDAYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTALPQTAYGDAYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120
QY 141 NHMGVDGAGSSVDYVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNVSLPDLDTTK 200
DB 121 NHMGVDGAGSSVDYVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNVSLPDLDTTK 180
QY 201 DVVKNEWYDMWGLSVNSYIDGLRIDTVKHVKQDFMPGYNKAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDMWGLSVNSYIDGLRIDTVKHVKQDFMPGYNKAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSSTLLGTTFVENHNDPR 320
DB 241 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSSTLLGTTFVENHNDPR 300
QY 321 FASYTNDIALAKNVAAFIILNDGIPPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPPIIYAGQEQHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420
QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 5
US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6022724o No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
```

```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/683,838A
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/600,908
/ FILING DATE: 13-FEB-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Green, Reza
/ REGISTRATION NUMBER: 38,475
/ REFERENCE/DOCKET NUMBER: 4394.204-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-867-0123
/ TELEFAX: 212-878-9655
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 478 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match          95.8%; Score 2572; DB 3; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGGSTTATCNTADQKYGCGTWQGIIDKLDYIQGMGFTAI 80
Db 1 ATPADWRSQSIYFLLTDRFARTDGGSTTATCNTADQKYGCGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTALPQTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
Db 61 WITPVTALPQTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120

QY 141 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNTVSLPDLDTTK 200
Db 121 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDWVGLSVNSYISDGLRIDTVKHVQKDFWPGYNKAAGVYVCIGEVLDGDPAYTC 260
Db 181 DVVKNEWYDWVGLSVNSYISDGLRIDTVKHVQKDFWPGYNKAAGVYVCIGEVLDGDPAYTC 240

QY 261 PYQNVMGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSDSTLLGTTFVENHDNPR 320
Db 241 PYQNVMGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSDSTLLGTTFVENHDNPR 300

QY 321 FASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGNDPANREATWLSGYPTDSELYK 380
Db 301 FASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
Db 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
Db 421 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 6
US-09-182-859-7
; Sequence 7, Application US/09182859
; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Biegard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/672,459
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30

/
/ TITLE OF INVENTION: Alpha-Amylase Mutants
/ FILE REFERENCE: 4796.204-US
/ CURRENT APPLICATION NUMBER: US/09/182,859
/ CURRENT FILING DATE: 1998-10-29
/ EARLIER APPLICATION NUMBER: 0515/96
/ EARLIER FILING DATE: 1996-04-30
/ EARLIER APPLICATION NUMBER: 0712/96
/ EARLIER FILING DATE: 1996-06-28
/ EARLIER APPLICATION NUMBER: 0775/96
/ EARLIER FILING DATE: 1996-07-11
/ EARLIER APPLICATION NUMBER: 1263/96
/ EARLIER FILING DATE: 1996-11-08
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 7
/ LENGTH: 478
/ TYPE: PRT
/ ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match          95.8%; Score 2572; DB 3; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGGSTTATCNTADQKYGCGTWQGIIDKLDYIQGMGFTAI 80
Db 1 ATPADWRSQSIYFLLTDRFARTDGGSTTATCNTADQKYGCGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTALPQTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
Db 61 WITPVTALPQTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120

QY 141 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNTVSLPDLDTTK 200
Db 121 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTVQEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDWVGLSVNSYISDGLRIDTVKHVQKDFWPGYNKAAGVYVCIGEVLDGDPAYTC 260
Db 181 DVVKNEWYDWVGLSVNSYISDGLRIDTVKHVQKDFWPGYNKAAGVYVCIGEVLDGDPAYTC 240

QY 261 PYQNVMGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSDSTLLGTTFVENHDNPR 320
Db 241 PYQNVMGVLNPIYYPLLNAPKSTSGSMDDLYNMINTVKSCDPSDSTLLGTTFVENHDNPR 300

QY 321 FASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGNDPANREATWLSGYPTDSELYK 380
Db 301 FASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
Db 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
Db 421 LSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 7
US-09-672-459-7
; Sequence 7, Application US/09672459
; Patent No. 6436888
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Biegard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/672,459
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
```

;; PRIOR APPLICATION NUMBER: 0712/96
;; PRIOR FILING DATE: 1996-06-28
;; PRIOR APPLICATION NUMBER: 0775/96
;; PRIOR FILING DATE: 1996-07-11
;; PRIOR APPLICATION NUMBER: 1263/96
;; PRIOR FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 7
;; LENGTH: 478
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-09-672-459-7

Query Match 95.8%; Score 2572; DB 4; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGTATCNCNTADQKCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRFARTDGTATCNCNTADQKCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTQAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 140
DB 61 WITPVTQAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 120

QY 141 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYKAAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYKAAAGVYCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLYNMINTVKS DCPDSTLLGTTFVENHDNPR 320
DB 241 PYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLYNMINTVKS DCPDSTLLGTTFVENHDNPR 300

QY 321 FASYTNDIALAKNVAAFIILNDGIPIIYAGQBOHYAGGNDPANREATWLSGYPTDSSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPIIYAGQBOHYAGGNDPANREATWLSGYPTDSSELYK 360

QY 381 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 8
US-09-636-252A-10
; Sequence 10, Application US/09636252A
; Patent No. 6440716
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/1F216-US2
; CURRENT APPLICATION NUMBER: US/09/636, 252A
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683, 838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 478
; TYPE: PRT
; ORGANISM: A. oryzae
US-09-636-252A-10

Query Match 95.8%; Score 2572; DB 4; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGTATCNCNTADQKCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRFARTDGTATCNCNTADQKCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTQAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 140
DB 61 WITPVTQAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 120

QY 141 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYKAAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYKAAAGVYCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLYNMINTVKS DCPDSTLLGTTFVENHDNPR 320
DB 241 PYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLYNMINTVKS DCPDSTLLGTTFVENHDNPR 300

QY 321 FASYTNDIALAKNVAAFIILNDGIPIIYAGQBOHYAGGNDPANREATWLSGYPTDSSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPIIYAGQBOHYAGGNDPANREATWLSGYPTDSSELYK 360

QY 381 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQOLTEVIGCTTVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 9
US-10-186-042-7
; Sequence 7, Application US/10186042
; Patent No. 6642044
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/10/186, 042
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/672, 459
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182, 859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-186-042-7

Query Match 95.8%; Score 2572; DB 4; Length 478;
Best Local Similarity 99.8%; Pred. No. 3.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSSQSIYFLLTDRFARTDGSSTATCNTADQKYGCGTWWQGIIDKLDYIQGMGTAI 80
Db 1 ATPADWRSSQSIYFLLTDRFARTDGSSTATCNTADQKYGCGTWWQGIIDKLDYIQGMGTAI 60
QY 81 WITPVTAQLPQTATAYGDAYHGTVQODIYSLNENYGTADDLKALSSALHERGMYLWVDVVA 140
Db 61 WITPVTAQLPQTATAYGDAYHGTVQODIYSLNENYGTADDLKALSSALHERGMYLWVDVVA 120
QY 141 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVSLPDLDTTK 200
Db 121 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVSLPDLDTTK 180
QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 260
Db 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVLNPIYIYPLNNAFKSTSGSMDDLYNNINTVKSDDPSTLLGTTFVENHNDNR 320
Db 241 PYQNVMDGVLNPIYIYPLNNAFKSTSGSMDDLYNNINTVKSDDPSTLLGTTFVENHNDNR 300
QY 321 FASYTNDIALAKNVAAFIILNDGPIIYAGQOHYAGGNDPANREATWLSGYPTDSELYK 380
Db 301 FASYTNDIALAKNVAAFIILNDGPIIYAGQOHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
Db 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420
QY 441 LSLGAGYTAGQQLTEVIGCTTVTVGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
Db 421 LSLGAGYTAGQQLTEVIGCTTVTVGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 10

US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306096/1993
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiser, Allen S.
; REGISTRATION NUMBER: 27,315
; REFERENCE/DOCKET NUMBER: 18335.009
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-339-715A-2
Query Match 93.3%; Score 2503.5; DB 2; Length 478;
Best Local Similarity 97.9%; Pred. No. 4e-214;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;
QY 21 ATPADWRSSQSIYFLLTDRFARTDGSSTATCNTADQKYGCGTWWQGIIDKLDYIQGMGTAI 80
Db 1 ATPADWRSSQSIYFLLTDRFARTDGSSTATCNTADQKYGCGTWWQGIIDKLDYIQGMGTAI 60
QY 81 WITPVTAQLPQTATAYGDAYHGTVQODIYSLNENYGTADDLKALSSALHERGMYLWVDVVA 140
Db 61 WITPVTAQLPQTATAYGDAYHGTVQODIYSLNENYGTADDLKALSSALHERGMYLWVDVVA 120
QY 141 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVSLPDLDTTK 200
Db 121 NHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTVQEDCWLGDNVSLPDLDTTK 180
QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 260
Db 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGVLNPIYIYPLNNAFKSTSGSMDDLYNNINTVKSDDPSTLLGTTFVENHNDNR 320
Db 241 PYQNVMDGVLNPIYIYPLNNAFKSTSGSMDDLYNNINTVKSDDPSTLLGTTFVENHNDNR 300
QY 321 FASYTNDIALAKNVAAFIILNDGPIIYAGQOHYAGGNDPANREATWLSGYPTDSELYK 380
Db 301 FASYTNDIALAKNVAAFIILNDGPIIYAGQOHYAGGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
Db 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 419
QY 441 LSLGAGYTAGQQLTEVIGCTTVTVGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498
Db 420 LSLGAGYTAGQQLTEVIGCTTVTVGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 477

RESULT 11

US-08-470-702-6
; Sequence 6, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

THIS PAGE BLANK (USPTO)

Run on: February 8, 2005, 18:42:07 ; Search time 214.39 Seconds
(without alignments)
756.564 Million cell updates/sec

Sequence: 1 MVAWSLFLYGLQVAAPALA.....LPRVLYPTEKLAGSKICSSS 498

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1373511

Maximum Match 100%
Listing first 45

Database : PUBLISHED_APPLICATIONS_AA:
1: /cqn2_6/ptodata/2/pubnaaa

2: /cqn2_6/ptodata/2/pubpaa/FCI_NEW_FOB.pcp:
3: /cqn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*

5: /cgn2_6/ptodata/2/pubpaa/pctus pubcomb.per:
6: /cgn2_6/ptodata/2/pubpaa/pctus pubcomb.per:

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pép:*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.per

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp

```
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp
15: /cgn2_6/ntodata/2/pubpaa/US10C_PUBCOMB.pcp
```

15: /cg2-6/ptodata/2/pubpaa/US10D_PUBCOMB.pep
16: /cg2-6/ptodata/2/pubpaa/US10D_PUBCOMB.pep

19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp:

20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.per:

Pred. No. is the number of results predicted by chance to

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

—

No.	Score	Match	Length	DB	ID	Description
1	2674	99.6	499	16	US-10-815-495-18	Sequence 18, Appl
2	2662	99.2	498	16	US-10-815-495-22	Sequence 22, Appl
3	2572	95.8	478	14	US-10-184-771-10	Sequence 10, Appl
4	2572	95.8	478	14	US-10-186-042-7	Sequence 7, Appl
5	2572	95.8	478	15	US-10-644-187-7	Sequence 7, Appl
6	2572	95.8	478	17	US-10-926-120-10	Sequence 10, Appl
7	2569	95.7	495	14	US-10-228-063-45	Sequence 45, Appl
8	2100	78.2	495	14	US-10-213-590-42	Sequence 42, Appl
9	1835	68.4	630	14	US-10-213-590-45	Sequence 45, Appl
10	1778	66.2	484	15	US-10-416-393-1	Sequence 1, Appl
11	1214.5	45.2	547	14	US-10-213-590-48	Sequence 48, Appl
12	1160	43.2	500	15	US-10-369-493-12469	Sequence 12469, A
13	1015.5	37.8	513	15	US-10-369-493-2223	Sequence 2223, Ap

QY	181	VEDCWLGDNTVSLPDLDTTKDVVKNEWYDWVGSLSVNSYIDGLRIDTVKHVKQDFWPGYN	240
Db	182	VEDCWLGDNTVSLPDLDTTKDVVKNEWYDWVGSLSVNSYIDGLRIDTVKHVKQDFWPGYN	241
QY	241	KAAGVYCIGEVLDGDPAYTCPQNVMDGVNLPIYYPLLNAPKSTSGSMDDLYNMINTVK	300
Db	242	KAAGVYCIGEVLDGDPAYTCPQNVMDGVNLPIYYPLLNAPKSTSGSMDDLYNMINTVK	301
QY	301	SDCPDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQHQHAGGND	360
Db	302	SDCPDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQHQHAGGND	361
QY	361	PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTYKNWPIYKDDTTIAMRKGT	420
Db	362	PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTYKNWPIYKDDTTIAMRKGT	421
QY	421	DGSIQIVTILSNKGASGDSYTLISLGAGYTAGQOLTVEIGCTTVTSGDGNVPVPMAGGLP	480
Db	422	DGSIQIVTILSNKGASGDSYTLISLGAGYTAGQOLTVEIGCTTVTSGDGNVPVPMAGGLP	481
QY	481	RVLYPTEKLAGSKICSSS	498
Db	482	RVLYPTEKLAGSKICSSS	499
RESULT 2			
US-10-815-495-22			
; Sequence 22, Application US/10815495			
; Publication No. US20040191864A1			
; GENERAL INFORMATION:			
; APPLICANT: Novozymes Biotech, Inc.			
; APPLICANT: Connolly, Mariah			
; APPLICANT: Brody, Howard			
; TITLE OF INVENTION: Methods For Producing Biological Substances In Enzyme-Deficient			
; FILE REFERENCE: Mutants Of Aspergillus Niger			
; FILE REFERENCE: 10345.200-US			
; CURRENT APPLICATION NUMBER: US/10/815,495			
; CURRENT FILING DATE: 2004-03-31			
; NUMBER OF SEQ ID NOS: 30			
; SOFTWARE: Patentin version 3.2			
; SEQ ID NO 22			
; LENGTH: 498			
; TYPE: PRT			
; ORGANISM: Aspergillus niger			
US-10-815-495-22			
Query Match 99.2%; Score 2662; DB 16; Length 498;			
Best Local Similarity 99.6%; Pred. No. 1.6e-211;			
Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	MVAWWSLFYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGTATCNTADQKYCGG	60
Db	2	MVAWWSLFYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGTATCNTADQKYCGG	61
QY	61	TWGGIIDLKDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDL	120
Db	62	TWGGIIDLKDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDL	121
QY	121	KALSSALHERGMVLMYVNVANHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYBDQTO	180
Db	122	KALSSALHERGMVLMYVNVANHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYBDQTO	181
QY	181	VEDCWLGDNTVSLPDLDTTKDVVKNEWYDWVGSLSVNSYIDGLRIDTVKHVKQDFWPGYN	240
Db	182	VEDCWLGDNTVSLPDLDTTKDVVKNEWYDWVGSLSVNSYIDGLRIDTVKHVKQDFWPGYN	241
QY	241	KAAGVYCIGEVLDGDPAYTCPQNVMDGVNLPIYYPLLNAPKSTSGSMDDLYNMINTVK	300
Db	242	KAAGVYCIGEVLDGDPAYTCPQNVMDGVNLPIYYPLLNAPKSTSGSMDDLYNMINTVK	301
QY	301	SDCPDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQHQHAGGND	360

Db	302	SDCPDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQHQHAGGND	361
QY	361	PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTYKNWPIYKDDTTIAMRKGT	420
Db	362	PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTYKNWPIYKDDTTIAMRKGT	421
QY	421	DGSIQIVTILSNKGASGDSYTLISLGAGYTAGQOLTVEIGCTTVTSGDGNVPVPMAGGLP	480
Db	422	DGSIQIVTILSNKGASGDSYTLISLGAGYTAGQOLTVEIGCTTVTSGDGNVPVPMAGGLP	481
QY	481	RVLYPTEKLAGSKIC	495
Db	482	RVLYPTEKLAGSKIC	496
RESULT 3			
US-10-184-771-10			
; Sequence 10, Application US/10184771			
; Publication No. US20030170769A1			
; GENERAL INFORMATION:			
; APPLICANT: Svendsen, Allan			
; APPLICANT: Bisgard-Frantzen, Henrik			
; APPLICANT: Borchert, Torben Vedel			
; TITLE OF INVENTION: Alpha-Amylase Mutants			
; FILE REFERENCE: 0776/1F216-US2			
; CURRENT APPLICATION NUMBER: US/10/184,771			
; CURRENT FILING DATE: 2002-06-28			
; PRIOR APPLICATION NUMBER: US/09/636,252			
; PRIOR FILING DATE: 2000-08-10			
; PRIOR APPLICATION NUMBER: 08/683,838			
; PRIOR FILING DATE: 1996-07-18			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 10			
; LENGTH: 478			
; TYPE: PRT			
; ORGANISM: A. oryzae			
US-10-184-771-10			
Query Match 95.8%; Score 2572; DB 14; Length 478;			
Best Local Similarity 99.8%; Pred. No. 4.1e-204;			
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	21	ATPADWRSQSIYFLLTDRFARTDGTATCNTADQKYCGGTWQGIIDKLDYIQGMGFTAI	80
Db	1	ATPADWRSQSIYFLLTDRFARTDGTATCNTADQKYCGGTWQGIIDKLDYIQGMGFTAI	60
QY	81	WITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDLKALSSALHERGMVLMYDVVA	140
Db	61	WITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDLKALSSALHERGMVLMYDVVA	120
QY	141	NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK	200
Db	121	NHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK	180
QY	201	DVVKNNEWYDWVGSLSVNSYIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC	260
Db	181	DVVKNNEWYDWVGSLSVNSYIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC	240
QY	261	PYQNVMDGVNLPIYYPLLNAPKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHNDNR	320
Db	241	PYQNVMDGVNLPIYYPLLNAPKSTSGSMDDLYNMINTVKSDCPDSTLLGTFVENHNDNR	300
QY	321	FASYNTDIALAKNVAAFIILNDGIPIIYAGQHQHAGGNDPANREATWLSGYPTDSELYK	380
Db	301	FASYNTDIALAKNVAAFIILNDGIPIIYAGQHQHAGGNDPANREATWLSGYPTDSELYK	360
QY	381	LIASANAIRNVAISKDTGFVTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT	440
Db	361	LIASANAIRNVAISKDTGFVTYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT	420
QY	441	LSLSGAGYTAGQOLTVEIGCTTVTSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS	498

Db 421 LSLSGAGYTAGQQLTEVIGCTTGTGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 4

US-10-186-042-7

; Sequence 7, Application US/10186042

; Publication No. US20030171236A1

; GENERAL INFORMATION:

; APPLICANT: Svendsen, Allan

; APPLICANT: Bisgard-Frantzen, Henrik

; TITLE OF INVENTION: Alpha-Amylase Mutants

; FILE REFERENCE: 4796.204-US

; CURRENT APPLICATION NUMBER: US/10/186,042

; CURRENT FILING DATE: 2002-06-28

; PRIOR APPLICATION NUMBER: US/09/672,459

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: 09/182,859

; PRIOR FILING DATE: 1998-10-29

; PRIOR APPLICATION NUMBER: 0515/96

; PRIOR FILING DATE: 1996-04-30

; PRIOR APPLICATION NUMBER: 0712/96

; PRIOR FILING DATE: 1996-06-28

; PRIOR APPLICATION NUMBER: 0775/96

; PRIOR FILING DATE: 1996-07-11

; PRIOR APPLICATION NUMBER: 1263/96

; PRIOR FILING DATE: 1996-11-08

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7

; LENGTH: 478

; TYPE: PRT

; ORGANISM: Bacillus licheniformis

US-10-186-042-7

Query Match 95.8%; Score 2572; DB 14; Length 478;

Best Local Similarity 99.8%; Pred. No. 4.1e-204;

Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGSATTATCNCADQKCYCGGTWQGIIDKLDYIQGMGFTAI 80

Db 1 ATPADWRSQSIYFLLTDRFARTDGSATTATCNCADQKCYCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTALPQTAYGDVHGYYQDDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 140

Db 61 WITPVTALPQTAYGDVHGYYQDDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 120

QY 141 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTVEDCWLGDNTVSLPDLDTTK 200

Db 121 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTVEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYKKAAGVYCIGEVLDGDPAYTC 260

Db 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYKKAAGVYCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDLLYNNINTVKSCDPSDSTLLGTFFVENHDNPR 320

Db 241 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDLLYNNINTVKSCDPSDSTLLGTFFVENHDNPR 300

QY 321 FASYTNDIALAKNVAFFIILNDGIPIIYAGQOHYAGGNDPANREATWLSGYPTDSSELYK 380

Db 301 FASYTNDIALAKNVAFFIILNDGIPIIYAGQOHYAGGNDPANREATWLSGYPTDSSELYK 360

QY 381 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIARMKGTGDSQIVTILSNKGASGDSYT 440

Db 361 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIARMKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQQLTEVIGCTTGTGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498

Db 421 LSLSGAGYTAGQQLTEVIGCTTGTGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 5

US-10-644-187-7

; Sequence 7, Application US/10644187

; Publication No. US20040048351A1

; GENERAL INFORMATION:

; APPLICANT: Svendsen, Allan

; APPLICANT: Bisgard-Frantzen, Henrik

; TITLE OF INVENTION: Alpha-Amylase Mutants

; FILE REFERENCE: 4796.204-US

; CURRENT APPLICATION NUMBER: US/10/644,187

; CURRENT FILING DATE: 2003-08-20

; PRIOR APPLICATION NUMBER: 09/182,859

; PRIOR FILING DATE: 1998-10-29

; PRIOR APPLICATION NUMBER: 0515/96

; PRIOR FILING DATE: 1996-04-30

; PRIOR APPLICATION NUMBER: 0712/96

; PRIOR FILING DATE: 1996-06-28

; PRIOR APPLICATION NUMBER: 0775/96

; PRIOR FILING DATE: 1996-07-11

; PRIOR APPLICATION NUMBER: 1263/96

; PRIOR FILING DATE: 1996-11-08

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7

; LENGTH: 478

; TYPE: PRT

; ORGANISM: Bacillus licheniformis

US-10-644-187-7

Query Match 95.8%; Score 2572; DB 15; Length 478;

Best Local Similarity 99.8%; Pred. No. 4.1e-204;

Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGSATTATCNCADQKCYCGGTWQGIIDKLDYIQGMGFTAI 80

Db 1 ATPADWRSQSIYFLLTDRFARTDGSATTATCNCADQKCYCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTALPQTAYGDVHGYYQDDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 140

Db 61 WITPVTALPQTAYGDVHGYYQDDIYSLNENYGTADDL KALSSALHERGMYLMDVVA 120

QY 141 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTVEDCWLGDNTVSLPDLDTTK 200

Db 121 NHMGYDAGSSVDYSVFKPFSSQDYFHPFCFQNYEDQTVEDCWLGDNTVSLPDLDTTK 180

QY 201 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYKKAAGVYCIGEVLDGDPAYTC 260

Db 181 DVVKNEWYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYKKAAGVYCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDLLYNNINTVKSCDPSDSTLLGTFFVENHDNPR 320

Db 241 PYQNVMDGVLNPIYYPLLNAPKSTSGSMDLLYNNINTVKSCDPSDSTLLGTFFVENHDNPR 300

QY 321 FASYTNDIALAKNVAFFIILNDGIPIIYAGQOHYAGGNDPANREATWLSGYPTDSSELYK 380

Db 301 FASYTNDIALAKNVAFFIILNDGIPIIYAGQOHYAGGNDPANREATWLSGYPTDSSELYK 360

QY 381 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIARMKGTGDSQIVTILSNKGASGDSYT 440

Db 361 LIASANAIRNVAISKDTGFTYKNWPIYKDDTTIARMKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQQLTEVIGCTTGTGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 498

Db 421 LSLSGAGYTAGQQLTEVIGCTTGTGSDGNVPPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 6

US-10-926-720-10

; Sequence 10, Application US/10926720

; Publication No. US20050019886A1

; GENERAL INFORMATION:

; APPLICANT: Svendsen, Allan

; Bisg+rd-Frantzen, Henrik

;/ Borchert, Torben Vedel
;/ TITLE OF INVENTION: -Amylase Mutants
;/ NUMBER OF SEQUENCES: 13
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Novo Nordisk of North America, Inc.
;/ STREET: 405 Lexington Avenue, 64th Floor
;/ CITY: New York
;/ STATE: New York
;/ COUNTRY: United States of America
;/ ZIP: 10174-6401
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: Patent In Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/10/926,720
;/ FILING DATE: 26-Aug-2004
;/ CLASSIFICATION: <Unknown>
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/600,908
;/ FILING DATE: 13-FEB-1996
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Green, Reza
;/ REGISTRATION NUMBER: 38,475
;/ REFERENCE/DOCKET NUMBER: 4394.204-US
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 212-867-0123
;/ TELEFAX: 212-878-9655
;/ INFORMATION FOR SEQ ID NO: 10:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 478 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: cDNA
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-926-720-10

Query Match 95.8%; Score 2572; DB 17; Length 478;
Best Local Similarity 99.8%; Pred. No. 4.1e-204;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGSSTATCNCADQKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRFARTDGSSTATCNCADQKYCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTAQLPQTAYGDYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTAQLPQTAYGDYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120

QY 141 NHMGYDVGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGNTVSLPDLDTTK 200
DB 121 NHMGYDVGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGNTVSLPDLDTTK 180

QY 201 DVVKNWYDVGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNWYDVGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYPIYPLNAPKSTSGSMDDLYNNINTVKSCDPDSTLLGTFTVENHDNPR 320
DB 241 PYQNVMDGVLNPIYPIYPLNAPKSTSGSMDDLYNNINTVKSCDPDSTLLGTFTVENHDNPR 300

QY 321 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQQLTEVIGCTTTVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQQLTEVIGCTTTVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

Db 421 LSLSGAGYTAGQQLTEVIGCTTTVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 7
US-10-228-063-45
; Sequence 45, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Aspergillus shiirousami
US-10-228-063-45

Query Match 95.7%; Score 2569; DB 14; Length 1095;
Best Local Similarity 99.8%; Pred. No. 2.3e-203;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADWRSQSIYFLLTDRFARTDGSSTATCNCADQKYCGGTWQGIIDKLDYIQGMGFTAI 80
DB 1 ATPADWRSQSIYFLLTDRFARTDGSSTATCNCADQKYCGGTWQGIIDKLDYIQGMGFTAI 60

QY 81 WITPVTAQLPQTAYGDYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 140
DB 61 WITPVTAQLPQTAYGDYHGYWQDIYSLNENYGTADDLKALSSALHERGMYLMDVVA 120

QY 141 NHMGYDVGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGNTVSLPDLDTTK 200
DB 121 NHMGYDVGAGSSVDYSVFKPFSSQDYFHPFCFIONYEDQTVEDCWLGNTVSLPDLDTTK 180

QY 201 DVVKNWYDVGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 260
DB 181 DVVKNWYDVGSLVSNYSIDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYPIYPLNAPKSTSGSMDDLYNNINTVKSCDPDSTLLGTFTVENHDNPR 320
DB 241 PYQNVMDGVLNPIYPIYPLNAPKSTSGSMDDLYNNINTVKSCDPDSTLLGTFTVENHDNPR 300

QY 321 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAAFIILNDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 420

QY 441 LSLSGAGYTAGQQLTEVIGCTTTVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGQQLTEVIGCTTTVTGVSDGNVPVPMAGGLPRVLYPTEKLAGSKICSSS 478

RESULT 8
US-10-213-990-42
; Sequence 42, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Busesey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72


```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Aspergillus
US-10-213-990-42

Query Match      78.2%; Score 2100; DB 14; Length 495;
Best Local Similarity 78.6%; Pred. No. 5e-165;
Matches 389; Conservative 39; Mismatches 65; Indels 2; Gaps 1;

QY 2 VAWNSFLYGLQVAAPALAAAPADWRSOSIYFLLTDRFARTDGSSTATCNTADQKVCQT 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IRWSSFFLSCL--AGTALAAATPAQWRSOSIYFLLTDRFARTDGSSTASCDTSARBYCGT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 WQGIIDKLDYIQGMGFTAIWITPVAQLPQTAYGDAYGHYQQDIYSLNENYGTADDLK 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 WQGIIEQLDYIQGMGFTAIWITPVTQLPQDTSEGTAYGHYQQDIYSVNSNYGTADDLK 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALSALHERGMVLMVDVVAHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTV 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 ALSALHHRGMVLMVDVVAHMGYAGAGSDYSVFENPFNSQTSFPLPLCFISYDQNTDV 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 182 EDCVLGNTVSLPDLDTTKVVKNEMWDMVGSLSVNSYIDGLRIDTVKHVQKDFWPGYNK 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 ENCWLGDNSVPLPDLDTTNPVQKIWNWVNSLSVNSYIDGLRIDTVKHVQSDFWPGFND 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 242 AAGVYCEIGELVDGDPAYTCPYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLYNMINTVKS 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 AAGVYCEIGELVDGDPAYTCPQEVLDGVLNPIYYPYLLKAFQSTSGSSSLYDMINTVKS 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 302 DCPDSTLLGTFFVENHNDNPRPASYNTNDIALAKNVAAPFIILNDGPIIYAGQEQHYAGGNDP 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 QCADSTLLGTFFVENHNDTFRFASYTKDMALAKNAAAFIIFSDGPIIYAGQEQHYSGGADP 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 362 ANREATWLSGYPTDSELKYLITASANAIRNYAISKDTGPVTKWNIYKDDTTIAMRKGTD 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 ANREAVWLSGYSTTSDLYKLITATANAIRSHAISKDPGYVTKYKNPIYKDTSTIAMRKGSD 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 422 GSQIVTILSNKGASGDSYTLISLGAGYTAGOOLTEVIGCTTVTVGSDGNVPVPMAGELPR 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 GAQIITVLSNLGASGSSYTLISLGGTYEAGQOLTENFSCCTTVTVGSDKKVPVSMASGLPR 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 482 VLYPTEKLAGSKICS 496
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 VFYPTAGLNGSTVCT 495
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-10-213-990-45
; Sequence 45, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storey, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Aspergillus
US-10-213-990-45

Query Match      68.4%; Score 1835; DB 14; Length 630;
Best Local Similarity 66.8%; Pred. No. 6e-143;
Matches 334; Conservative 59; Mismatches 101; Indels 6; Gaps 2;

```


Dd	421	FRKRGREGRIIMVLSTQGSNGAYTIRMAN-GFQPSVAVARDVFSCRTWTVDNMGSLRLDM	479
Qy	476	AGGLPRVLYPTEKLAKSKIC	495
Dd	480	DKGEPRVLRPEALMRGSLC	499
 RESULT 13			
US-10-369-493-2223			
; Sequence 2223, Application US/10369493			
; Publication No. US20030233675A1			
; GENERAL INFORMATION:			
; APPLICANT: Cao, Yongwei			
; APPLICANT: Hinkle, Gregory J.			
; APPLICANT: Slater, Steven C.			
; APPLICANT: Goldman, Barry S.			
; APPLICANT: Chen, Xianfeng			
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF			
; FILE REFERENCE: 38-10(52052)B			
; CURRENT APPLICATION NUMBER: US/10/369,493			
; PRIOR FILING DATE: 2003-02-28			
; PRIOR APPLICATION NUMBER: US 60/360,039			
; PRIOR FILING DATE: 2002-02-21			
; NUMBER OF SEQ ID NOS: 47374			
; SEQ ID NO 2223			
; LENGTH: 513			
; TYPE: PRT			
; ORGANISM: Schizosaccharomyces pombe			
US-10-369-493-2223			
 Query Match 37.8%; Score 1015.5; DB 15; Length 513;			
Best Local Similarity 41.4%; Pred. No. 3e-75;			
Matches 204; Conservative 84; Mismatches 170; Indels 35; Gaps 12;			
Qy	5	WSL-FLYGLQVAAPALAATPADWRSSIIFFLALTDRFAFDGSGTTATCNTADOKYCGETWQ	63
Dd	9	WTLUVLSFLLVLPANALDKHGWRKSISLUTDRPASTN---PKPCNPEDREBYCGGNWR	65
Qy	64	GIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHYGWQDIYSLNENYGTADDLKAL	123
Dd	66	GIIDKLDYIQGMGFTAIWISPIIKNEGRTKYEAYHGYWPQDLYLNFPHFGTEODLLDL	125
Qy	124	SSALHRGRMYLVMDVANHWYDYGAGSSVDYVSFPKPFSSQDYVHFPCFINVEDOTQVED	183
Dd	126	ADALHDGRMYLVMDTVVNHWG--SSDPNTIDYGIYRPPFNQS SHYPWCPTIEQ-DKPLSLEQ	183
Qy	184	CWLGDNTVSLPDLDTTKDVVKNEMWDVGSLYSNSISIGLRDTHVKHVQKDFWPGYNKAA	243
Dd	184	CWGTGSDMTLPDIDTENPOIETFLYNFIHQKQFKIDGLRVDAHKVRRTFWPGFCESA	243
Qy	244	GVCICGEVLDDGPAYTCPYQNVMGVLNYPI-----YYPLLN-AFKSTSGSMDDLNNM	295
Dd	244	GVCYQGEWVTGQADLFCEQEYMDGLHNFPVOGVAAESVIPLNDRALRKTAIAM-----	297
Qy	296	INTVKSDCDSPDTLLGTFVENHNPNRPASVTNDIALAKNVAAFIILNDGPIIYAGQEQHY	355
Dd	298	NLVAHHCKDSITLLGLFLESQAPRLAALNNDYTVLKNAMTLNLMSDGIPIVYFQGQMF	356
Qy	356	AGGNDPANREATWLSGYPTDSLEYLKLIASANAIRNAVISKDTG--FVYKKNWPIYKDDTT	413
Dd	357	NGSHDPVNRPALMOOQYNTDGELYQYTSKVNKIRDLINSEDEGIYIRSITHAIMIGHV	416
Qy	414	IAMRKGTGDSQIVTILSNKGASDYSYTLSSLGAGYTAGQQLTEVTCGGTTVTVSGDNV-P	472
Dd	417	MWKYG---PVITFITNGAVDKVEYLKIOWPGS-----ETMDILLTCLTLEV--EGEVNR	465
Qy	473	VPMAGGLPRVLYP	485
Dd	466	TSIKKGEPKILYP	478

RESULT 14

US-10-369-493-22679
; Sequence 22679, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22679
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-22679

Query Match 33.5%; Score 898.5; DB 15; Length 564;
Best Local Similarity 38.8%; Pred. No. 1.6e-65;
Matches 188; Conservative 85; Mismatches 190; Indels 21; Gaps 7;

QY 18 ALAATPADWRSSQSIYELLTDRFARTDGGTTATCNTADOKYCGTWOGLIIDLVDYIQMGF 77
DB 19 ANAGSNAERKRIYLIQIUTDRFVDDGGTDNDPCDPANQYCGTWOGLIIDLVDYIQMGF 78
QY 78 TAIWITPVTAQLP-QTTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMVLMV 136
DB 79 NAIWISPIDKNIEGDIDGAGYAHGYWNTDYESLNEHFGCTEDDLVSLITAARKAGIWMVL 138
QY 137 DVVANHGVDGAGSSVDYVFKPFSSQDYFHPFCFI--QNYEDQTOVEDCWLGDNTVSLP 194
DB 139 DSIVNSMALAPPLADADYSSLSNPFNKESYFHPYCLIDWDITDNETNVMDCW-QDSGYLLA 197
QY 195 DLDITTKDVKNWYDWGSLVSNYSIDGLRIDTVKHVKDFWPGYNKAAGVYCIQGVLDG 254
DB 198 DLDVSSDSVSLSDHFKSLISKYDFDGLRIDAVQNMNYTFPFDVDTGVTVSGVFSY 257
QY 255 DPAYTCYQNVMDGLNYPYIYPLLNAPKSTSGSMDLLYNMINTVKSDCP--DSTLLGTF 312
DB 258 DPDTMCSYMLPGVTNYFLQLYINFSFATGAGFTLIPTYQEVMAASCKYDSTLMLTF 317
QY 313 VENHDNPRFASYTNDIALAKVAAPIILNDGIPYIYAGOEQHYAGGNDPANREATWLSGY 372
DB 318 IENHDLRFPYPTSDSQIMGALSFVLIDWGTPISFYGOEQFGNGEDPANRPAWLTGY 377
QY 373 PTSELYKLASANAIRNYALSKTGFVTVKNWPIYKDDTTIAMRGTGTD----GSOIVT 427
DB 378 DQSNFYVTVTKMWAFKRFVITQPDWVT-----STYOSIENVHVVQKQNDLV 428
QY 428 ILSNKGASGDSYTLSSLAGYTAGOQLTEVIGCTTWTGSGDGNVPVPMAGGLPRVLYPTE 487
DB 429 MENWNGVT--NNLTIEYETNYTANEVVSDVFGHRTLTVGADKTLTASMTNGYELIMYPHS 487
QY 488 KLAG 491
DB 488 KMSG 491

RESULT 15
US-10-369-493-22723
; Sequence 22723, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:28:34 ; Search time 57.6834 Seconds
(without alignments)
830.671 Million cell updates/sec

Title: US-10-820-200-2
Perfect score: 2684
Sequence: 1 MWAWSLFLYGLQVAAPALLA.....LPRVLPTEKLAGSKICSSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2684	100.0	499	1 ALAS1	alpha-amylase (EC
2	2675	99.7	499	2 JS0663	alpha-amylase (EC
3	2674	99.6	499	1 ALAS3	alpha-amylase (EC
4	2674	99.6	499	2 B48305	alpha-amylase (EC
5	2673	99.6	499	2 JN0588	alpha-amylase (EC
6	2662	99.2	498	2 A48305	alpha-amylase (EC
7	2656	99.0	499	2 JT0466	alpha-amylase (EC
8	2503.5	93.3	478	2 JK0201	alpha-amylase (EC
9	1780	66.3	484	1 A35282	alpha-amylase (EC
10	1629	60.7	624	1 JC4510	pullulanase (EC 3.
11	1527	56.9	507	2 S33921	alpha-amylase (EC
12	1449	54.0	512	2 S06115	alpha-amylase (EC
13	1445.5	53.9	512	2 S23355	alpha-amylase (EC
14	1397.5	52.1	494	1 ALBYAP	alpha-amylase (EC
15	1264	47.1	631	2 S72270	alpha-amylase (EC
16	1015.5	37.8	513	2 T38770	alpha-amylase a pr
17	963.5	35.9	625	2 T41603	alpha-amylase - fi
18	920.5	34.3	491	2 T38448	probable alpha-am
19	899	33.5	581	2 S62505	probable alpha-am
20	898.5	33.5	584	2 T41503	alpha-amylase - fi
21	898	33.5	478	2 T40860	probable alpha-am
22	737	27.5	482	2 S31478	alpha-amylase (EC
23	670	25.0	1196	2 T39539	beta-amylase (EC 3
24	653.5	24.3	774	2 T39539	alpha-amylase homo
25	481.5	17.9	642	2 A11827	cyclomaltodextrin
26	456.5	17.0	712	1 ALBSG1	cyclomaltodextrin
27	454.5	16.9	712	1 ALBSG3	cyclomaltodextrin
28	444.5	16.6	713	1 ALBSG7	cyclomaltodextrin
29	440.5	16.4	528	1 ALBSK	alpha-amylase (EC

RESULT 1
ALAS1
alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
N;Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
C;Species: Aspergillus oryzae
C;Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 09-Jul-2004
C;Accession: S04548; A33214; JS0240; A91930; A93767; A10627
R;Wiseel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A;Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon o
A;Reference number: S04548; MUID:89237897; PMID:2785629
A;Accession: S04548
A;Molecule type: DNA
A;Residues: 1-499 <WIR>
A;Cross-references: UNIPROT:P10529; EMBL:X12725; NID:G2430; PIDN:CAA31218.1; PID:G295921
A;Genetics: AMY1
A;Accession: A33214
A;Molecule type: mRNA
A;Residues: 1-499 <W12>
R;Genes, M.J.; Dove, M.J.; Seligy, V.L.
Gene 79, 107-117, 1989
A;Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A;Reference number: JS0240; MUID:89378767; PMID:2789162
A;Accession: JS0240
A;Molecule type: DNA
A;Residues: 1-499 <GEN>
A;Genetics: AMY2
A;Note: the authors refer to this as isozyme II
R;Isemura, S.; Ikenaka, T.
J. Biochem. 74, 1-10, 1973
A;Reference number: A91930; MUID:74001521; PMID:4733850
A;Accession: A91930
A;Molecule type: protein
A;Residues: 206-225 <ISE>
R;Narita, K.
Proc Jpn. Acad. 51, 285-290, 1975
A;Reference number: A93767
A;Accession: A93767
A;Molecule type: protein
A;Residues: 434-443, 446-447, 'O', 449-458, 'GTTV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>
R;Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A;Title: Structure and possible catalytic residues of Taka-amylase A.
A;Reference number: A37454; MUID:84212370; PMID:6609921
A;Contents: annotation; X-ray crystallography, 3.0 angstroms
R;Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
submitted to the Brookhaven Protein Data Bank, August 1992
A;Reference number: A51548; PDB:6TAA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
C;Comment: One atom of calcium per molecule is essential for activity.
C;Genetics: <AMY1>

ALIGNMENTS

30	433	16.1	710	2	S63598
31	428.5	16.0	655	1	ALBEG
32	428	15.9	703	1	ALBSX1
33	424.5	15.8	919	2	S28179
34	418	15.6	713	2	A58800
35	418	15.6	714	1	ALBSGR
36	414.5	15.4	713	2	S09196
37	412	15.4	717	1	S28784
38	410	15.3	483	2	G75392
39	407	15.2	704	2	I39805
40	406.5	15.1	718	1	ALBSGC
41	405.5	15.1	718	1	ALBSG6
42	405	15.1	713	1	ALBSXR
43	399.5	14.9	718	1	ALBSMX
44	384.5	14.3	711	1	ALBSXF
45	376.5	14.0	1104	2	A60999

A:Gene: amy1
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Genetics: <AMY2>
A:Gene: amy2; Amy11
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium, carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase 1 #status experimental
F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 2684; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 7e-185;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAWNSFLYGLQVAAPALAAPADWRSSQSIYFLLTDRFARTDGGSTATCNDADQKYCGG 60
DB 2 MVAWNSFLYGLQVAAPALAAPADWRSSQSIYFLLTDRFARTDGGSTATCNDADQKYCGG 61

QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDL 121

QY 121 KALSSALHERGMVLMVDVNVANHMGYDAGSSVDYSYFKPFSSQDYFHPFCFIQNYEDQTQ 180
DB 122 KALSSALHERGMVLMVDVNVANHMGYDAGSSVDYSYFKPFSSQDYFHPFCFIQNYEDQTQ 181

QY 181 VEDCWLGDNVSLPDLDTTKDVVKNEWYDWVGLSVNSYISDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCWLGDNVSLPDLDTTKDVVKNEWYDWVGLSVNSYISDGLRIDTVKHVKQDFWPGYN 241

QY 241 KAAGVYICIGVLDGDPAYTCYQNVMDGVNLPIYVPLNNAFKSTSGSMDLLYNNMINTVK 300
DB 242 KAAGVYICIGVLDGDPAYTCYQNVMDGVNLPIYVPLNNAFKSTSGSMDLLYNNMINTVK 301

QY 301 SDPCDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPPIYAGQEQHYAGND 360
DB 302 SDPCDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPPIYAGQEQHYAGND 361

QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 421

QY 421 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQOLTEVIGCTTGTGSDGNVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQOLTEVIGCTTGTGSDGNVPVPMAGGLP 481

QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499

RESULT 2
JS0663
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: JS0663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Bioc. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shiroyamii and its expression
A:Reference number: JS0663; MUID:92323146; PMID:1368777
A:Molecule type: mRNA
A:Residues: 1-499 <SHI>
C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <ALP>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.7%; Score 2675; DB 2; Length 499;
Best Local Similarity 99.8%; Pred. No. 3.1e-184;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVAWNSFLYGLQVAAPALAAPADWRSSQSIYFLLTDRFARTDGGSTATCNDADQKYCGG 60
DB 2 MVAWNSFLYGLQVAAPALAAPADWRSSQSIYFLLTDRFARTDGGSTATCNDADQKYCGG 61

QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDDIYSLNENYGTADDL 121

QY 121 KALSSALHERGMVLMVDVNVANHMGYDAGSSVDYSYFKPFSSQDYFHPFCFIQNYEDQTQ 180
DB 122 KALSSALHERGMVLMVDVNVANHMGYDAGSSVDYSYFKPFSSQDYFHPFCFIQNYEDQTQ 181

QY 181 VEDCWLGDNVSLPDLDTTKDVVKNEWYDWVGLSVNSYISDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCWLGDNVSLPDLDTTKDVVKNEWYDWVGLSVNSYISDGLRIDTVKHVKQDFWPGYN 241

QY 241 KAAGVYICIGVLDGDPAYTCYQNVMDGVNLPIYVPLNNAFKSTSGSMDLLYNNMINTVK 300
DB 242 KAAGVYICIGVLDGDPAYTCYQNVMDGVNLPIYVPLNNAFKSTSGSMDLLYNNMINTVK 301

QY 301 SDPCDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPPIYAGQEQHYAGND 360
DB 302 SDPCDSTLLGTFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPPIYAGQEQHYAGND 361

QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 421

QY 421 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQOLTEVIGCTTGTGSDGNVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQOLTEVIGCTTGTGSDGNVPVPMAGGLP 481

QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499

RESULT 3
ALAS3
alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
A:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S04549; A33215; A44713
R:Wirsel, S.; Lechmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon organization
A:Reference number: S04549; MUID:89237897; PMID:2785629
A:Molecule type: DNA
A:Residues: 1-499 <WIR>
A:Cross-references: UNIPROT:Q96TH4; EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
A:Accession: A33215
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-499 <W12>
A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing

A;Reference number: JS0240; MUID:89378767; PMID:2789162

A;Accession: A44713

A;Molecule type: DNA

A;Residues: 1-499 <GEN>

A;Note: the authors refer to this as isozyme I

J. Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A;Title: Structure and possible catalytic residues of Taka-amylose A.

A;Reference number: A37454; MUID:84212370; PMID:6609921

A;Contents: annotation; X-ray crystallography, 3.0 angstroms

C;Comment: One atom of calcium per molecule is essential for activity.

C;Genetics:

A;Gene: amy3; AmyI

A;Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C;Function:

A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A;Pathway: glycogen/starch degradation

C;Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-499/Product: alpha-amylose 3 #status experimental

F;194-321/Domain: alpha-amylose core homology <AMY>

F;51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental

F;142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F;218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F;231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 99.6%; Score 2674; DB 1; Length 499;

Best Local Similarity 99.6%; Pred. No. 3.7e-184;

Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVAWMSLFYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATNTADQKYCGG 60

DB 2 MVAWMSLFYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATNTADQKYCGG 61

QY 61 TWGGIIDLKLDYIOGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL 120

DB 62 TWGGIIDLKLDYIOGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL 121

QY 121 KALSSALHERGMVLMVDVVAHMGYDAGSSVDYSYVFKPSSQDYFPPFCFLQNYEDQIQ 180

DB 122 KALSSALHERGMVLMVDVVAHMGYDAGSSVDYSYVFKPSSQDYFPPFCFLQNYEDQIQ 181

QY 181 VEDCWLGDNTVSLPDLDTTKDVVKNWYDVGSLVSNYSIDGLRIDTVKHVQKDFWPGYN 240

DB 182 VEDCWLGDNTVSLPDLDTTKDVVKNWYDVGSLVSNYSIDGLRIDTVKHVQKDFWPGYN 241

QY 241 KAAGVYCI GEVLGDGPAYTCYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLNNMINTVK 300

DB 242 KAAGVYCI GEVLGDGPAYTCYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLNNMINTVK 301

QY 301 SDPCDSTLLGTFFVENHNDPRFASYNIDIALAKNVAAPILNDGIPPIYAGQEOHYAGGND 360

DB 302 SDPCDSTLLGTFFVENHNDPRFASYNIDIALAKNVAAPILNDGIPPIYAGQEOHYAGGND 361

QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTKNWPYIKDDTTIANKRGT 420

DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTKNWPYIKDDTTIANKRGT 421

QY 421 DGSQIVTILSNKAGSDSYTSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLP 480

DB 422 DGSQIVTILSNKAGSDSYTSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLP 481

QY 481 RVLVPTTEKLAGSKICSSS 498

DB 482 RVLVPTTEKLAGSKICSSS 499

RESULT 4

B48305

alpha-amylose (EC 3.2.1.1) B precursor - Aspergillus awamori

C;Species: Aspergillus awamori

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C;Accession: B48305

R;Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;

Curr. Genet. 17, 203-212, 1990

A;Title: Cloning, characterization, and expression of two alpha-amylose genes from Aspergillus

A;Reference number: A48305; MUID:90254827; PMID:2340591

A;Accession: B48305

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-499 <KOR>

A;Cross-references: UNIPROT:Q02906

C;Function:

A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A;Pathway: glycogen/starch degradation

C;Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

F;194-321/Domain: alpha-amylose core homology <AMY>

Query Match 99.6%; Score 2674; DB 2; Length 499;

Best Local Similarity 99.6%; Pred. No. 3.7e-184;

Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVAWMSLFYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATNTADQKYCGG 60

DB 2 MVAWMSLFYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATNTADQKYCGG 61

QY 61 TWGGIIDLKLDYIOGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL 120

DB 62 TWGGIIDLKLDYIOGMGFTAIWITPVTAQLPQTAYGDAYHGYWQDIYSLNENYGTADDL 121

QY 121 KALSSALHERGMVLMVDVVAHMGYDAGSSVDYSYVFKPSSQDYFPPFCFLQNYEDQIQ 180

DB 122 KALSSALHERGMVLMVDVVAHMGYDAGSSVDYSYVFKPSSQDYFPPFCFLQNYEDQIQ 181

QY 181 VEDCWLGDNTVSLPDLDTTKDVVKNWYDVGSLVSNYSIDGLRIDTVKHVQKDFWPGYN 240

DB 182 VEDCWLGDNTVSLPDLDTTKDVVKNWYDVGSLVSNYSIDGLRIDTVKHVQKDFWPGYN 241

QY 241 KAAGVYCI GEVLGDGPAYTCYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLNNMINTVK 300

DB 242 KAAGVYCI GEVLGDGPAYTCYQNVMDGVLNPIYYPYLLNAPKSTSGSMDLLNNMINTVK 301

QY 301 SDPCDSTLLGTFFVENHNDPRFASYNIDIALAKNVAAPILNDGIPPIYAGQEOHYAGGND 360

DB 302 SDPCDSTLLGTFFVENHNDPRFASYNIDIALAKNVAAPILNDGIPPIYAGQEOHYAGGND 361

QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTKNWPYIKDDTTIANKRGT 420

DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFVTKNWPYIKDDTTIANKRGT 421

QY 421 DGSQIVTILSNKAGSDSYTSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLP 480

DB 422 DGSQIVTILSNKAGSDSYTSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLP 481

QY 481 RVLVPTTEKLAGSKICSSS 498

DB 482 RVLVPTTEKLAGSKICSSS 499

RESULT 5

JN0588

alpha-amylose (EC 3.2.1.1) precursor - Aspergillus oryzae

N;Alternate names: Taka-amylose A

C;Species: Aspergillus oryzae

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

R;Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Uda, S.

Gene 84, 319-327, 1989

A;Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylose A: evidence for mu

A;Reference number: JN0588; MUID:90128276; PMID:2612911

A;Accession: JN0588

A;Molecule type: mRNA

A;Residues: 1-499 <TSU>

A;Cross-references: UNIPROT:Q96TH4

C:Comment: The alpha amylases are encoded by multigene family.

C:Genes: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <MNT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Aen) (covalent) #status predicted

```
Query Match          99.6%; Score 2673; DB 2; Length 499;
Best Local Similarity 99.6%; Pred. No. 4.3e-184;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVAMWSFLYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGSSTTATCNTADQKCGG 60
DB 2 MVAMWSFLYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGSSTTATCNTADQKCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMVLMVDVVAHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTO 180
DB 122 KALSSALHERGMVLMVDVVAHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTO 181
QY 181 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKGVQKDFWPGYN 240
DB 182 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKGVQKDFWPGYN 241
QY 241 KAAGVYCIGEVLGDGPAYTCPCQNVMDGVNLPIIYPLLNAFKSTSGSMDLLYNNINTVK 300
DB 242 KAAGVYCIGEVLGDGPAYTCPCQNVMDGVNLPIIYPLLNAFKSTSGSMDLLYNNINTVK 301
QY 301 SDCPOSTLLGTFFVENHNDPRFASYNIDIALAKNVAAFIILNDGIPPIIYAGQSHVAGGND 360
DB 302 SDCPOSTLLGTFFVENHNDPRFASYNIDIALAKNVAAFIILNDGIPPIIYAGQSHVAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGSAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGSAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499
```

RESULT 6

A48305
alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: A48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
A:Accession: A48305
A:Reference number: A48305; MUID:90254827; PMID:2340591
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
A:Cross-references: UNIPROT:Q02905
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

```
Query Match          99.2%; Score 2662; DB 2; Length 498;
Best Local Similarity 99.6%; Pred. No. 2.7e-183;
Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVAMWSFLYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGSSTTATCNTADQKCGG 60
DB 2 MVAMWSFLYGLQVAAPALAAATPADWRSQSIYFLLTDRFARTDGSSTTATCNTADQKCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHGWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMVLMVDVVAHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTO 180
DB 122 KALSSALHERGMVLMVDVVAHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQTO 181
QY 181 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKGVQKDFWPGYN 240
DB 182 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKGVQKDFWPGYN 241
QY 241 KAAGVYCIGEVLGDGPAYTCPCQNVMDGVNLPIIYPLLNAFKSTSGSMDLLYNNINTVK 300
DB 242 KAAGVYCIGEVLGDGPAYTCPCQNVMDGVNLPIIYPLLNAFKSTSGSMDLLYNNINTVK 301
QY 301 SDCPOSTLLGTFFVENHNDPRFASYNIDIALAKNVAAFIILNDGIPPIIYAGQSHVAGGND 360
DB 302 SDCPOSTLLGTFFVENHNDPRFASYNIDIALAKNVAAFIILNDGIPPIIYAGQSHVAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGSAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGSAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKIC 495
DB 482 RVLYPTEKLAGSKIC 496
```

RESULT 7

JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
A:Accession: JT0466
R:Itada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <RAD>
A:Cross-references: UNIPROT:P10529
C:Comment: See also PIR:JK0201 and PIR:JS0240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <MNT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Aen) (covalent) #status predicted

F;231,251,318/Active site: His, Glu, Asp #status predicted

```
Query Match 99.0%; Score 2656; DB 2; Length 499;
Best Local Similarity 99.4%; Pred. No. 7,2e-183;
Matches 495; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 MVAWWSLFLYGLVQVAAPALAAATPADRSQSIIYELLTDREARTDGSATTATCNDADQKCYCGG 60
DB 2 MVAWWSLFLYGLVQVAAPALAAATPADRSQSIIYELLTDREARTDGSATTATCNDADQKCYCGG 61

QY 61 TWQGIIDKLDYIQGMFTAIWITPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMFTAIWITPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDL 121

QY 121 KALSSALHERGMYLMDVVANHHGVDGAGSSVDYSVFKPFSSQDYPFHPFCFIQNYEDQIQ 180
DB 122 KALSSALHERGMYLMDVVANHHGVDGAGSSVDYSVFKPFSSQDYPFHPFCFIQNYEDQIQ 181

QY 181 VEDCWLGDNTVSLPDLDTTKVKNWYDMVGLSVNSYSDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEYCWLGDNVSLPDLDTTKVKNWYDMVGLSVNSYSDGLRIDTVKHVKQDFWPGYN 241

QY 241 KAAGVYICIGVLDGDPAYTCFYQNVMDGVLNPIYPLNAPKSTSGSMDLLNMINTVK 300
DB 242 KAAGVYICIGVLDGDPAYTCFYQNVMDGVLNPIYPLNAPKSTSGSMDLLNMINTVK 301

QY 301 SDPCDSTLLGTFFVENHNDNPFASYTNDIALAKNVAAPILNDGIPPIYAGQEQHYAGGND 360
DB 302 SDPCDSTLLGTFFVENHNDNPFASYTNDIALAKNVAAPILNDGIPPIYAGQEQHYAGGND 361

QY 361 PANREATWLSGYTDSSELYKLIASANAIRYAIKSDTGFVYKQWPIYKDDTTIAMRKG 420
DB 362 PANREATWLSGYTDSSELYKLIASANAIRYAIKSDTGFVYKQWPIYKDDTTIAMRKG 421

QY 421 DGSQIVTILSNKGASGDSYTLSSGAGYTAGQQLTEVIGCTTGVGSDGNVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLSSGAGYTAGQQLTEVIGCTTGVGSDGNVPMAGGLP 481

QY 481 RVLPTPEKLAGSKICSSS 498
DB 482 RVLPTPEKLAGSKICSSS 499

RESULT 8
JK0201
alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase, taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C:Accession: JK0201
R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A:Title: The complete amino acid sequence of Taka-amylase A.
A:Reference number: JK0201
A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOD>
A:Cross-references: UNIPROT:P10529
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:JS0240.
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;173-300/Domain: alpha-amylase core homology <AMY>
F;210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 93.3%; Score 2503.5; DB 2; Length 478;
Best Local Similarity 97.9%; Pred. No. 5,9e-172;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;
```

```
QY 21 ATPADRSQSIIYELLTDREARTDGSATTATCNDADQKCYCGGTWQGIIDKLDYIQGMFTAI 80
DB 1 ATPADRSQSIIYELLTDREARTDGSATTATCNDADQKCYCGGTWQGIIDKLDYIQGMFTAI 60

QY 81 WITPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVVA 140
DB 61 WITPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVVA 120

QY 141 NHMGYDGASSVDYSVFKPFSSQDYPFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDGASSVDYSVFKPFSSQDYPFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTK 180

QY 201 DVYKNEWYDMVGLSVNSYSDGLRIDTVKHVKQDFWPGYNKAAGVYICIGVLDGDPAYTC 260
DB 181 DVYKNEWYDMVGLSVNSYSDGLRIDTVKHVKQDFWPGYNKAAGVYICIGVLDGDPAYTC 240

QY 261 PYQNVMDGVLNPIYPLNAPKSTSGSMDLLNMINTVKSDCPDSTLLGTFFVENHNDNPR 320
DB 241 PYQNVMDGVLNPIYPLNAPKSTSGSMDLLNMINTVKSDCPDSTLLGTFFVENHNDNPR 300

QY 321 FASYTNDIALAKNVAAPILNDGIPPIYAGQEQHYAGGNDPANREATWLSGYTDSSELYK 380
DB 301 FASYTNDIALAKNVAAPILNDGIPPIYAGQEQHYAGGNDPANREATWLSGYTDSSELYK 360

QY 381 LIASANAIRYAIKSDTGFVYKQWPIYKDDTTIAMRKGTDGSIQIVTILSNKGASGDSY 440
DB 361 LIASANAIRYAIKSDTGFVYKQWPIYKDDTTIAMRKGTDGSIQIVTILSNKGASGDSY 419

QY 441 LSI-SGAGYTAGQQLTEVIGCTTGVGSDGNVPMAGGLPRVLYPTEKLAGSKICSSS 498
DB 420 LSI-SGAGYTAGQQLTEVIGCTTGVGSDGNVPMAGGLPRVLYPTEKLAGSKICSSS 477

RESULT 9
A35282
alpha-amylase (EC 3.2.1.1) - Aspergillus niger
C:Species: Aspergillus niger
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35282
R:Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; Peter
Biochemistry 29, 6244-6249, 1990
A:Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom res
A:Reference number: A35282; PMID:91002514; PMID:2207069
A:Accession: A35282
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trar
A:Molecule type: mRNA
A:Residues: 1-484 <BOE>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;173-300/Domain: alpha-amylase core homology <AMY>

Query Match 66.3%; Score 1780; DB 1; Length 484;
Best Local Similarity 67.2%; Pred. No. 5,1e-120;
Matches 319; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

QY 24 ADWRQSIIYELLTDREARTDGSATTATCNDADQKCYCGGTWQGIIDKLDYIQGMFTAIWIT 83
DB 4 AEWRQSIIYELLTDREARTDGSATTATCNDADQKCYCGGTWQGIIDKLDYIQGMFTAIWIS 63

QY 84 PVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERGMYLMDVVANHM 143
DB 64 PITEQLPQTADGEAVHYGQQDIYDVNSNFGTADDLKLSDALHARGMYLMDVVVPM 123

QY 144 GYDGAGSSVDYSVFKPFSSQDYPFHPFCFIQNYEDQTVEDCWLGDNTVSLPDLDTTKDV 203
DB 124 GYAGGNDVDYSVDFPFDSSSYFHPYCLITDMDNLTVMQDCWEGDITVSLPDLTTTAV 183

QY 204 KNEWYDMVGLSVNSYSDGLRIDTVKHVKQDFWPGYNKAAGVYICIGVLDGDPAYTCPIQ 263
DB 184 KNEWYDMVGLSVNSYSDGLRIDTVKHVKQDFWPGYNKAAGVYICIGVLDGDPAYTCPIQ 243
```

Db 184 RIWTDWADLVNSVSDGLRDSVLEVEPDPFFPGYQBAAGVYCVGEVDNGNPALDCPYQ 243

QY 264 NYMDGVLNYPYIYPLNAPKSTSGSMDDLYNMINTVKSDCPDSTLLGTFTVENHNDNPRFAS 323

Db 244 KVLGDVLNYPYTWQLLYAFESSGSGISNLNMIKSVASDCSDPTLLGNFIENHNDNPRFAS 303

QY 324 YNDIALAKNVAAFILNDGIPPIIYAGQHQHYAGNDPANRATWLSGPTDSELYKILIA 383

Db 304 YTSDSQAKNVLSPFLSDGIPVIVAGEHQHYSGGKVPYNREATWLSGPDYTSAEELYTWIA 363

QY 384 SANARNVAISKDTCGFVTYKNWPIYKDDTTIAMRKGTGDSQIIVTILSNKAGSDSYTSL 443

Db 364 TTNARKKLAISADSAYITANDAFYTDSTIAMRKGTSGSQIVTILSNKSGSSSYTTL 423

QY 444 SGAGYTAGQOLTEVIGCTTIVTSGDGNVPMAGGLPRVLYPTEKLAGSKICSSS 498

Db 424 SSGSYTSGTKLEAYTCTSVTVDSSGDIIPVPMASGLPRVLLPASVVDSSSLCGGS 478

RESULT 10

JC4510

N:pullulanase (EC 3.2.1.41) precursor - yeast (lipomyces kononenkoae)

N:Alternate names: LKAI protein; raw starch-degrading amylase

N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)

C:Species: lipomyces kononenkoae

C>Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004

C:Accession: JC4510; PC4116

R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.

Gene 166, 65-71, 1995

A:Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a Lipid

A:Reference number: JC4510; MUID:96105202; PMID:8529895

A:Accession: JC4510

A:Molecule type: mRNA

A:Residues: 1-624 <STE>

A:Cross-references: UNIPROT:Q01117; GB:U030376; NID:g1173536; PIDN:AAC49622.1; PID:g11735

A:Experimental source: strain IGC4052B

A:Accession: PC4116

A:Molecule type: protein

A:Residues: 29-44 <ST2>

A:Experimental source: IGC4052B

C:Genetics:

A:Gene: LKAI

C:Function:

A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages

A:Pathway: glycogen/starch degradation

C:Superfamily: lipomyces alpha-amylase; alpha-amylase core homology; glucoamylase starch

C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-624/Product: alpha-amylase #status predicted <WAT>

F:48-141/Domain: glucoamylase starch-binding domain homology <SBD>

F:320-447/Domain: alpha-amylase core homology <AMY>

F:177-185,297-311,387-430,587-622/Diulfide bonds: #status predicted

F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 60.7%; Score 1629; DB 1; Length 624;

Best Local Similarity 60.6%; Pred. No. 5e-109;

Matches 292; Conservative 77; Mismatches 113; Indels 0; Gaps 0;

QY 15 AAPALAAATPADWRSOSIYFLLTDRFARTDGSSTATCNTADQKYGCGTWOIHKLDYIQG 74

Db 142 SASVPTGTAAANRGRSIYQVWTRFARTDGSITVSCDVTRVYCGSGYGIINMLDIYQG 201

QY 75 MGFATWIITPVTAQLPQTAYGDAYHYGQQDIYSLNENYGTADDLKALSSALHREGMYL 134

Db 202 MGFATWIISPIVENIPDDTGYGYAHGYWMMKIDIFALNTNFGGADDLIALATELHNRGMYL 261

QY 135 MYDVVANHMGYDAGSSVDYVFKPSSQDYVFPFPFCIONVEDQTOVEDCWLGDNVTSLP 194

Db 262 MWDIVNHFAPSGNRADVISEYFPYSSQDYFHSFCWITDYDYNQTNVEECWLGDDSVPLV 321

QY 195 DLDTTKDVVKNEMYDWVGLSVSNYSIDGLRIDTVKHQKDFWPGYNKAAGVYCI GEVLDG 254

Db 322 DVNTQLDITVKSEYQSWVKQLIANYSIDGLRIDTVKHQKDFWAPQBAAGIYTVGEVFDG 381

QY 255 DPAYTCPYQNVMDGVLNYPYIYPLNAPKSTSGSMDDLYNMINTVKSDCPDSTLLGTFTVE 314

Db 382 DPSYTCPYQENLDGVLNYPYVYVVSFAFORVGGSSISLVDIMDTLKSECIDFTLLGSPLE 441

QY 315 NHDNPRFASYNNDIALAKNVAAFILNDGIPPIIYAGQHQHYAGNDPANRATWLSGYPT 374

Db 442 NQDNPRFPYSYSDSLIKNAIAFTILSDGIPPIIYAGQHQGLNGNDPNREALWPTGYST 501

QY 375 DSELYKILIASANARNVAISKDTCGFVTYKNWPIYKDDTTIAMRKGTGDSQIIVTILSNKA 434

Db 502 TSTFEYIASLNQIRNHAIIYDDTYLTYYQNWIIYSDSTIIAMRKGTGNQIITVLSNLGS 561

QY 435 SGDSYTLISLGSAGYTAGQOLTEVIGCTTIVTSGDGNVPMAGGLPRVLYPTEKLAGSKI 494

Db 562 SSGSYTTLISLTNGYTAGSSVVEILTCTAVTVDLNGLAVPMMSGGLPRVYFPESQLVSGSI 621

QY 495 CS 496

Db 622 CS 623

RESULT 11

S33921

N:alpha-amylase (EC 3.2.1.1) SWA2 precursor - yeast (Schwanniomyces occidentalis)

N:Alternate names: alpha-1,4 glucanohydrolase

C:Species: Schwanniomyces occidentalis

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

C:Accession: S33921

R:Claros, M.G.; Abarca, D.; Fernandez-Lobato, M.; Jimenez, A.

Curr. Genet. 24, 75-83, 1993

A:Title: Molecular structure of the SWA2 gene encoding an AMY1-related alpha-amylase from

A:Reference number: S33921; MUID:93365041; PMID:8358835

A:Accession: S33921

A:Molecule type: DNA

A:Residues: 1-507 <CLA>

A:Cross-references: UNIPROT:Q08806; EMBL:X73497; NID:g396561; PIDN:CAAS1912.1; PID:g39656

C:Genetics:

A:Gene: SWA2

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-507/Product: alpha-amylase #status predicted <WAT>

F:205-332/Domain: alpha-amylase core homology <AMY>

F:134,229/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 56.9%; Score 1527; DB 2; Length 507;

Best Local Similarity 57.7%; Pred. No. 7.8e-102;

Matches 286; Conservative 74; Mismatches 130; Indels 6; Gaps 3;

QY 1 MVAMWSLFYGLQVAAPALAAATPADWRSOSIYFLLTDRFARTDGSSTATCNTADQKYGCG 60

Db 17 LVASKPIFLSKRDAGSAAAA----WRSESIYQLVTRFARTDGSSTATCNTGDRVYCGG 72

QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDAYHYGQQDIYSLNENYGTADDL 120

Db 73 TFQGIIDKLDYIQGMGFTAIWISPVVEIQIPDDTGYGYAHGYWMMKIDIVAINSNFGTADDL 132

QY 121 KALSSALHREGMYLMDVVVANHMGYDAGSSVDYVFKPSSQDYVFPFPFCIONVEDQTOQ 180

Db 133 KNLSELHKKRNKMLMVDIVTNHYANWGAGSSVAYSNYNPFNQQSFDHYCLLTNTDDQIN 192

QY 181 VEDCWLGDNVTSLPDLDTTKDVVKNEMYDWVGLSVSNYSIDGLRIDTVKHQKDFWPGYN 240

Db 193 VEDCWEGDNVTSLPDLRTEDSDVSSIFNLWVAELVSNYSIDGLRIDSAKHVDESFPSPQ 252

QY 241 KAAGVYCI GEVLGDGPATVTCYQNVMDGVLNYPYIYPLNAPKSTSGSMDDLYNMINTVK 300

Db 253 SAAGVYLLGEVYDGDPAVTCYQNTMGSVTNYPLYPMLRFPFGQTSNSVDELNAMISSLE 312

[illegible]

A:Reference number: S00064: MUID:87276512: PMID:3497057

A:Accession number: S000604, NCBI:67276314, PMID:3437037
A:Molecule type: DNA
A:Residues: 1-494 <INFO>
A:Cross-references: UNIPROT:P21567; EMBL:X05791; NID:G4847; PIDN:CAA29233.1; PID:G4848
A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in having 368-Thr
C:Genetics:
A:Gene: ALP1
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; extracellular protein; glycoprotein; glycosidase; hydrolase; metallo
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-494/Product: alpha-amylase #status predicted <MAT>
F:200-327/Domain: alpha-amylase core homology <AMY>
F:57-151,267-310,462-493/Disulfide bonds: #status predicted
F:148,189,202,237/Binding site: calcium (Asn, Gln, Asp, His) #status predicted
F:224/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:233,257,324/Active site: Asp, Glu, Asp #status predicted

Query Match	52.1%;	Score 1397.5;	DB 1;	Length 494;
Best Local Similarity	52.5%;	Pred. No. 1.5e-92;		
Matches 264;	Conservative	83;	Mismatches 129;	Indels 27;
				Gaps 6;

Qy	12	LQVAAPALAA-----TPAD-WRSQSIIYFLLTDRPARTDGGSTATCNTA	53
Db	1	MOISKAALLASLAALVYAQPVTLPKRETNADKWRSQSIYQIVTRFARTDGDTSASCNTE	60
Qy	54	DOKYCGGTWOGIIDKLDIYIOGMGFTAIWITPVTAQLPOTTAYGDAYHYGWODIYSLSNEN	113
Db	61	DRLYCGGSFOGIKKLDIYIKONGFTAIIWISPVENIPDNNTAYGYAHGYMKWIYKINEN	120
Qy	114	YGTDADLKALSALHERGMYLMDVVVANHMGVDGAGSDVSVEFKPFSSODYTHFPFCFIQ	173
Db	121	FGTADDLKSIAQELHDDRMLLMVDIVTNHYGSDGSGDSIDYSEYTPENDQKYFHNYCLIS	180
Qy	174	NIYEDQTQVEDCWLGNNTYSLPDLDITTKOVKNWEYDWGSLVSNYSIDGLRIDITVGHVKQ	233
Db	181	NYDDQAQVQSCWEGDSSVALPOLRTEDSDVASVFNSMWKDPEGVNYSIDGLRIIDSAKHVDQ	240
Qy	234	DFWPGYNKAAGYCICEVLGDGPATCTCPYQNMDGVLMIPIYYPLLNAPKSTSGMSDDL	293
Db	241	GFFPDFVSASGVYSGVEVFGQDPATCTCPYQNVYICGVSNYPXYPTTFPKTTDSSSEL	300
Qy	294	NMINTVKSDCPSTLLTGTFVENHDNPRASTYNDIALAKXVAARIILNDGPIPIYAGOEQ	353
Db	301	QMISVASSCSOPTLLTFVENHDNERFASTSPQSJISNAIAFVLVGDPGPIVIYGOEQ	360
Qy	354	HVAGNGNDPANREATULSGPYPTDSELYKLITAGANIRYAIISKOTGFVYIKNWPIYKDDTT	413
Db	361	GLSGKSDPNREALWSLGVNKESDYKLIKANARAANAUVQDSYATSQLSVFNNDHV	420
Qy	414	IAMRKDTGSGIQIVTILSKNGASGDSYTLISLSGAGYTAGQQLTVEIGCTTVTVGSDGNVPV	473
Db	421	IATKRGS----VVSFNNILGSGSS-DVTISNTYSGCEDLVEVLTCSTVSGSSD--LOV	473

Qy 474 PMAGGLPRVLYPTEKLAGSKICS 496
: ||| : : : |||
Db 474 SIQGQPQIFVPAK--YASDICS 494

RESULT 15

S72270
alpha-amylase (EC 3.2.1.1) precursor - *Cryptococcus* sp. (strain CS2)
C:Species: *Cryptococcus* sp.
A:Variety: strain CS2
C:Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: S72270
R:Refuji, H.; Chino, M.; Kato, M.; Iimura, Y.
Biochem. J. 318, 989-996, 1996
A:Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus*
A:Reference number: S72270; MUID:96433120; PMID:8836148

A:Accession: S72270

A:Accession: G72270
A:Molecule type: DNA
A:Residues: 1-631 <IEF>
A:Cross-references: UNIPROT:Q92394; EMBL:D83540; NID:G1595852; PIDN:BAAL2010.1
A:Experimental source: Strain S-2
C:Genetics:
A:Gene: amy-CS2
A:Introns: 289/2; 326/2
C:Function:
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-631/Product: alpha-amylase #status predicted <WAT>
F:206-335/Domain: alpha-amylase core homology <AMY>

Query Match 47.1%; Score 1264; DB 2; Length 631;
Best Local Similarity 49.2%; Pred. No. 8e-83;
Matches 251; Conservative 84; Mismatches 153; Indels 2

Qy	9	LYGLQVAAPALAA--TPADWRSSQSIYFLLTRDPARTDGS7TATCN7ADQ--KYCGGTWQG	64
Db	7	LAGALLASGLVAGLSPAEWRSSQSIYQVVTDRFALDNGCNSPSCQSGSLNLYCNGTFFAG	66
Qy	65	IIDKLDYIQMGFTAIWITPVTAQLP-----QTTAYGDAYHCYQOODIYSLNENYGTADD	119
Db	67	IIDKLDYIQNGFTAIWISPVVVKINDGSGPNGYTPDGSAYHGYWAQDIYEINPHFGGAG	126
Qy	120	LKALSSALHERGMYLWMDVVANHWGY-----DGAGSSVDYSVFKPSSQDYVHPFCEI	172
Db	127	LTDLSNALHSRGMYLWMDVVVNHMAYYCTDGGCGPGSNVYGSTPTPNSBSYFHPFCEI	186
Qy	173	QNYEDQTVQVEDCWLGDNTVSLPDLDTTKDVVKNEWYDMVGSIVSNYSIDGLRIDTVKHVQ	232
Db	187	-DYNRTSILDCWEGDEIVPLVDLRTEDSDVQSIENSMISNLQIYINDGLAIDSLQQSG	245
Qy	233	KDFWPGYNKAA-GVYCTGEVLDDGDPAYTCPQNV-MDGVLNYPYIYPLLNAPKSTSGSMD	290
Db	246	SFFPFGFQAAGMYMVEGVENGSPSYVCPYQAGMPGVLYNPMFPFYITNAFPTSGSMS	305
Qy	291	DLYNWINTVSKDCPDSTLLGTFVFNHNDNPREASYTNDIALAKNVAAFIILNDGIPITYAG	350
Db	306	QLAGISAMOSDCSDTLLGSLFLENQDNPRFPSTQSLDTRAQNAFTMLQDGIPIYYG	365
Qy	351	QEQRHAGGNPANREATWLS-GYPTDSLEYKLIIASANAIRNYAISKDTGFVYQKWPIYK	409
Db	366	QEHLSSGVPLNREALWTSGGYDTSSELYEMIITVNQLRTLAIKONGGFVYKIQVPYT	425
Qy	410	DDTTIAMRKGTGDSQIVTILSNKAGSDSYTLSLGA--GYTAAQOLTEVICGTTTVVGS	467
Db	426	DSNHIIVTRKNGSGYQIVGVYTNVGSAGSSLSISSETSGFQASEPVMVDVLSCTLYHTGT	485
Qy	468	DGNVPVPMAGGLPRVILPTEKLAGSKITCSS	497
Db	486	DGSLFTMTGGLPRVFNATLAESSLCTT	515

Search completed: February 8, 2005, 18:43:01
Job time : 63.6834 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:25:28 ; Search time 292.263 Seconds
(without alignments)
872.556 Million cell updates/sec

Title: US-10-820-200-2

Perfect score: 2684

Sequence: 1 MWAWSLFLYGLQVAPALA.....LPRVLPTEKLAGSKICSS 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2684	100.0	498	2 Q76CT3	Q76ct3 aspergillus
2	2684	100.0	499	1 AMYA ASPOR	Pi0529 aspergillus
3	2680	99.9	499	2 Q76L99	P76199 aspergillus
4	2675	99.7	499	1 AMY ASPSH	P30292 aspergillus
5	2674	99.6	499	1 AMYB ASPAW	Q02906 aspergillus
6	2674	99.6	499	2 Q96TH4	Q96th4 aspergillus
7	2674	99.6	499	2 Q7LV45	Q7lv45 aspergillus
8	2662	99.2	498	1 AMYA ASPAW	Q02905 aspergillus
9	1826	68.0	490	2 Q9UV07	Q9uv07 emericella
10	1810	67.4	640	2 O13296	O13296 aspergillus
11	1791	66.7	634	2 Q76L96	Q76l96 aspergillus
12	1778	66.2	484	1 AMYA ASPNG	P56271 aspergillus
13	1725.5	64.3	623	2 Q9UV09	Q9uv09 emericella
14	1662	61.9	647	2 Q6YF33	Q6yf33 lipomyces s
15	1629	60.7	624	1 AMY1 LIPKO	Q01117 lipomyces k
16	1527	56.9	507	1 AMY2 DBBOC	Q08806 debaryomyce
17	1449	54.0	512	1 AMY1 DBBOC	P19269 debaryomyce
18	1397.5	52.1	494	1 AMY1_SACFI	P21567 saccharomyc
19	1272.5	47.4	492	2 Q7SDJ6	Q7sdj6 neurospora
20	1264	47.1	631	2 Q92394	Q92394 cryptococcu
21	1056.5	39.4	533	2 Q7S4K0	Q7s4k0 neurospora
22	1015.5	37.8	513	1 AMY3 SCHPO	O14154 schizosacch
23	963.5	35.9	625	2 Q74922	Q74922 schizosacch
24	920.5	34.3	491	2 O13996	O13996 schizosacch
25	899	33.5	581	1 AMY1 SCHPO	Q09840 schizosacch
26	898.5	33.5	564	1 AMY4 SCHPO	Q9y7e9 schizosacch
27	898	33.5	478	1 YQ29 SCHPO	O10427 schizosacch
28	895	33.3	499	2 Q8J1E4	Q8j1e4 lipomyces k
29	737	27.5	482	2 Q60U51	Q60u51 thermactin
30	670	25.0	1196	1 AMYB PAEPO	P21543 paenibacill
31	653.5	24.3	774	1 AMY2 SCHPO	O42918 schizosacch

RESULT 1

ID	Q76CT3	PRELIMINARY;	PRT;	498 AA.
AC	Q76CT3;			
DT	05-JUL-2004 (TReMBLrel. 27, Created)			
DT	05-JUL-2004 (TReMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TReMBLrel. 27, Last annotation update)			
DE	Alpha-amylase.			
GN	Name=amyA;			
OS	Aspergillus kawachi (Aspergillus awamori var. kawachi).			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=40384;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Ito K.;			
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.			
DR	EMBL; AB109452; BAD01051.1; -.			
DR	GO; GO:0004556; F:alpha-amylase activity; IEA.			
DR	GO; GO:0005975; P:carbohydrate metabolism; IEA.			
DR	InterPro; IPR006047; Alpha amyl cat.			
DR	InterPro; IPR006589; Alp_aml cat sub.			
DR	Pfam; PF006046; Glyco_hydro_13.			
DR	PRINTS; PR00110; ALPHAAMYLASE.			
DR	SMART; SM00642; Amy; 1.			
SQ	SEQUENCE 498 AA; 54679 MW; 8407BBFC7F4E3FC0 CRC64;			
Query Match	100.0%;	Score 2684;	DB 2;	Length 498;
Best Local Similarity	100.0%;	Pred. No. 2e-178;		
Matches 498;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1 MWAWSLFLYGLQVAPALAATPADWRQSIIYFLLTDRFARTDGSSTATCNTADQKCGG 60			
Db	1 MWAWSLFLYGLQVAPALAATPADWRQSIIYFLLTDRFARTDGSSTATCNTADQKCGG 60			
Qy	61 TWQGIIDKLDYIQGMGFTAIWITPTVTAQPTTAYGDAYHGYWQDDIYSLNENYGTADDL 120			
Db	61 TWQGIIDKLDYIQGMGFTAIWITPTVTAQPTTAYGDAYHGYWQDDIYSLNENYGTADDL 120			
Qy	121 KALSSALHERGMYLMVDVVANHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQIQ 180			
Db	121 KALSSALHERGMYLMVDVVANHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQIQ 180			
Qy	181 VEDCHLGDNTVSLPDLTDKDVVKNEWDVWCSLSVNSYIDGLRIDTVDKHKQDFWPGYN 240			
Db	181 VEDCHLGDNTVSLPDLTDKDVVKNEWDVWCSLSVNSYIDGLRIDTVDKHKQDFWPGYN 240			
Qy	241 KAAGVYCI GEVLDDGPAYTCPYQNVMDGLNPIYYPILLNAPKSTSGSMDDLYNNINTVK 300			
Db	241 KAAGVYCI GEVLDDGPAYTCPYQNVMDGLNPIYYPILLNAPKSTSGSMDDLYNNINTVK 300			
Qy	301 SDCPDSTLLGTFTVENHNDNPRFASNTNDIALAKNVAAFIILNDGIPITIVAGQEQHYAGGND 360			

32	481.5	17.9	642	2 Q820C9	Q820c9 anabaena sp
33	479.5	17.9	642	2 Q8RMG0	Q8rmg0 nostoc sp.
34	461.5	17.2	877	2 Q87FT5	Q87fts vibrio para
35	456.5	17.0	713	1 CDGT_BACSO	P05618 bacillus sp
36	454.5	16.9	712	1 CDGT_BACS3	P09121 bacillus sp
37	453	16.9	1798	2 Q9KZ11	Q9kz11 streptomyce
38	450	16.8	719	1 AMYM_BACST	P19531 bacillus st
39	447.5	16.7	712	2 Q6S3E3	Q6s3e3 bacillus sp
40	447.5	16.7	713	2 Q9F5W3	Q9f5w3 bacillus ci
41	444.5	16.6	713	1 CDGT_BAC11	P30921 bacillus ci
42	440.5	16.4	528	1 AMY_BACCI	P08137 bacillus ci
43	436	16.2	1806	2 Q82AS4	Q82as4 streptomyce
44	433	16.1	710	1 CDGT_THETU	P26827 thermoanaer
45	430.5	16.0	524	2 Q8R900	Q8r900 thermoanaer

ALIGNMENTS

```
Db 301 SDCPSTLGTVEVHNDRFASYNIDIALAKNVAFIILNDGIPIIYAGQHQHYAGND 360
Qy 361 PANREATWLSGYPTSELYKLTASANAIRNYAISKDTGFTVTKNWPYIKDDTTIAMRGT 420
Db 361 PANREATWLSGYPTSELYKLTASANAIRNYAISKDTGFTVTKNWPYIKDDTTIAMRGT 420
Qy 421 DGSQIVTILSNKGAGSDSYTSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVVPVWAGGLP 480
Db 421 DGSQIVTILSNKGAGSDSYTSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVVPVWAGGLP 480
Qy 481 RVLYPTEKLAGSKICSS 498
Db 481 RVLYPTEKLAGSKICSS 498

RESULT 2
ID AMYA ASPOR STANDARD; PRT; 499 AA.
AC P10529; P11763; Q00250;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase A precursor (BC 3.2.1.1) (Taka-amylase A) (TAA) (1,4-
alpha-D-glucan glucanohydrolase).
GN Name=AMY1;
GN and
GN Name=AMY2;
GN and
GN Name=AMY3;
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 63303;
RX MEDLINE=89237897; PubMed=2785629;
RA Wiesel S., Lachmund A., Wildhardt G., Rutkowski E.;
RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
intron-exon organization.";
RL Mol. Microbiol. 3:3-14(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89378767; PubMed=2789162; DOI=10.1016/0378-1119(89)90096-6;
RA Genes M.J., Dove M.J., Seligy V.L.;
RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
containing eight introns.";
RL Gene 79:107-117(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene of
Aspergillus oryzae.";
RL Agric. Biol. Chem. 53:593-599(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128276; PubMed=2612911; DOI=10.1016/0378-1119(89)90506-4;
RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A., Udaoka S.;
RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
evidence for multiple related genes.";
RL Gene 84:319-327(1989).
RN [5]
RP SEQUENCE OF 22-499.
RA Toda H., Kondo K., Narita K.;
RT "The complete amino acid sequence of Taka-amylase A.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
RN [6]
RP SEQUENCE OF 206-225.
RX MEDLINE=74001521; PubMed=4733850;
RA Isemura S., Ikenaka T.;
RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
A with trypsin and chymotrypsin.";
```

```
RL J. Biochem. 74:1-10(1973).
RN [7]
RP SEQUENCE OF 433-499.
RA Narita K.;
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=80227691; PubMed=6156152;
RA Matsuura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
Toda H., Narita K., Kakudo M.;
RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
3-A resolution.";
RL J. Biochem. 87:1555-1558(1980).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=84212370; PubMed=6609921;
RA Matsuura Y., Kusunoki M., Harada W., Kakudo M.;
RT "Structure and possible catalytic residues of Taka-amylase A.";
RL J. Biochem. 95:697-702(1984).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
RX MEDLINE=97428212; PubMed=9283074; DOI=10.1021/bi970539i;
RA Brzowski A.M., Davies G.J.;
RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
inhibitor acarbose at 2.0-A resolution.";
RL Biochemistry 36:10837-10845(1997).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
at high concentrations.
CC -!- SUBUNIT: Monomer.
CC -!- BIOTECHNOLOGY: Used in the brewing industry to increase the
fermentability of beer worts (including those made from unmalted
cereals), in the starch industry to make high maltose and high DE
syrops (starch saccharification), in the alcohol industry to
reduce fermentation time, in the cereal food industry for flour
supplementation and improvement of chilled and frozen dough, and
in the forestry industry for low-temperature modification of
starch. Sold under the name Fungamyl by Novozymes.
CC -!- MISCELLANEOUS: The sequence of AMY1 and AMY2 is shown.
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
RN [11]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
RN [12]
DR EMBL; X12725; CAA31218.1; -
DR EMBL; X12726; CAA31219.1; -
DR EMBL; X12727; CAA31220.1; -
DR EMBL; D00434; BAA00336.1; -
DR EMBL; M33218; AAA32708.1; -
DR PIR; JK0201; JK0201.
DR PIR; JT0466; JT0466.
DR PIR; S04548; ALAS1.
DR PDB; 2TAA; X-ray; A=22-499.
DR PDB; 6TAA; X-ray; @=22-499.
DR PDB; 7TAA; X-ray; @=22-499.
DR GlycoSuiteDB; p10529; -
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR006047; Alp_amyl_cat.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PRO0110; ALPHAAMYLASE.
DR SMART; SM00642; Amyv; 1.
DR 3D-structure; Calcium-binding; Carbohydrate metabolism;
KW Direct protein sequencing; Glycoprotein; Glycosidase; Hydrolase;
KW Multigene family; Signal.
FT SIGNAL 1 21
CHAIN 22 499 Alpha-amylase A.
```



```
Query Match      99.9%; Score 2680; DB 2; Length 499;
Best Local Similarity 99.8%; Pred. No. 3.9e-178;
Matches 497; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAWWSFLYGLQVAAPALAAATPADWRSQSIIYFLLTDRFARTDGGSTTATCMTADQKCGG 60
DB 2 MVAWWSFLYGLQVAAPALAAATPADWRSQSIIYFLLTDRFARTDGGSTTATCMTADQKCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDVHGYWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDVHGYWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMYLMVDVVAHMGYDGGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQQT 180
DB 122 KALSSALHERGMYLMVDVVAHMGYDGGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQQT 181
QY 181 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 241
QY 241 KAAGVYICIGEVLDGDPAYTCPYQNVMDGVNLPIYYPYLLNAPFKSTSGSMDDLYNNINTVK 300
DB 242 KAAGVYICIGEVLDGDPAYTCPYQNVMDGVNLPIYYPYLLNAPFKSTSGSMDDLYNNINTVK 301
QY 301 SDPCDSTLLGTFVFNHNDNPRFASYNNDIALAKNVAAFIILNDGIPYIYAGQEQHYAGGND 360
DB 302 SDPCDSTLLGTFVFNHNDNPRFASYNNDIALAKNVAAFIILNDGIPYIYAGQEQHYAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVTKNMPYIKDDTTIAMRGT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVTKNMPYIKDDTTIAMRGT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499

RESULT 4
AMY ASPSH
ID - AMY ASPSH .STANDARD; PRT; 499 AA.
AC P30292;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN Name=AMY;
OS Aspergillus shirousami.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillius.
OX NCBI_TaxID=5070;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92323146; PubMed=1368777;
RA Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT expression in Saccharomyces cerevisiae."
RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC -/- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -/- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -/- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10461; BAA01255.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR006589; Alp_aml_cat_sub.
DR InterPro; IPR006047; Alpha_aml_cat.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Amy; 1.
KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolase; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT ACT_SITE 142 142
FT METAL 183 183
FT METAL 196 196
FT METAL 227 227
FT METAL 231 231
FT METAL 251 251
FT DISULFID 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
SQ SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match      99.7%; Score 2675; DB 1; Length 499;
Best Local Similarity 99.8%; Pred. No. 8.6e-178;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVAWWSFLYGLQVAAPALAAATPADWRSQSIIYFLLTDRFARTDGGSTTATCMTADQKCGG 60
DB 2 MVAWWSFLYGLQVAAPALAAATPADWRSQSIIYFLLTDRFARTDGGSTTATCMTADQKCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDVHGYWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPQTAYGDVHGYWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMYLMVDVVAHMGYDGGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQQT 180
DB 122 KALSSALHERGMYLMVDVVAHMGYDGGSSVDYSVFKPFSSQDYFHPFCFTQNYEDQQT 181
QY 181 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWGSLVSNYSIDGLRIDTVKHVKQDFWPGYN 241
QY 241 KAAGVYICIGEVLDGDPAYTCPYQNVMDGVNLPIYYPYLLNAPFKSTSGSMDDLYNNINTVK 300
DB 242 KAAGVYICIGEVLDGDPAYTCPYQNVMDGVNLPIYYPYLLNAPFKSTSGSMDDLYNNINTVK 301
QY 301 SDPCDSTLLGTFVFNHNDNPRFASYNNDIALAKNVAAFIILNDGIPYIYAGQEQHYAGGND 360
DB 302 SDPCDSTLLGTFVFNHNDNPRFASYNNDIALAKNVAAFIILNDGIPYIYAGQEQHYAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVTKNMPYIKDDTTIAMRGT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVTKNMPYIKDDTTIAMRGT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGAGYTAGQQLTEVIGCTTIVTSGDGNVVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499
```

entites requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).


```

RESULT 5
AMBY ASPAW STANDARD; PRT; 499 AA.
ID Q02906;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylose B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN Name=AMYB;
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L., Kodama K.H.,
RA Royer T.J., Thompson S.A., Ward M., Wilson L.J., Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylose genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -! CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -! COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations (By similarity).
CC -! SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X52756; CAA36967.1; -.
DR PIR; B48305; B48305.
DR HSP; P10529; 7TAA.
DR InterPro; IPR006589; Alp_aml_cat_sub.
DR InterPro; IPR006047; Alpha_aml_cat.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylose; 1.
DR PRINTS; PR00110; ALPHAAMYLAASE.
DR SMART; SM00642; Amy; 1.
KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
KW Hydrolase; Multigene family; Signal.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 499 Alpha-amylose B.
FT ACT_SITE 227 227 Nucleophile (By similarity).
FT ACT_SITE 251 251 Proton donor (By similarity).
FT ACT_SITE 318 318 By similarity.
FT METAL 142 142 Calcium 1 (By similarity).
FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
FT similarity).
FT METAL 196 196 Calcium 1 (By similarity).
FT METAL 227 227 Calcium 2 (By similarity).
FT METAL 231 231 Calcium 1 (via carbonyl oxygen) (By
FT similarity).
FT METAL 251 251 Calcium 2 (By similarity).
FT DISULFID 51 59 By similarity.
FT DISULFID 171 185 By similarity.
FT DISULFID 261 304 By similarity.
FT DISULFID 461 496 By similarity.
FT CARBOHYD 218 218 N-linked (GLCNAC...) (Potential).
SQ SEQUENCE 499 AA; 54921 MW; 740896B11BGC1A8A CRC64;
Query Match 99.6%; Score 2674; DB 1; Length 499;
Best Local Similarity 99.6%; Pred. No. 1e-177; Indels 0; Gaps 0;
Matches 496; Conservative 0; Mismatches 2;

```

```

QY 1 MVAMWSLFLYGLQVAAPALAAATPADWRSOSIYFLLTDRFARTDGTATCNCATDQKCGG 60
DB 2 MVAMWSLFLYGLQVAAPALAAATPADWRSOSIYFLLTDRFARTDGTATCNCATDQKCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPTVTAQLPQTATAYGDAYHGYWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPTVTAQLPQTATAYGDAYHGYWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMYLMVYVNVANHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTO 180
DB 122 KALSSALHERGMYLMVYVNVANHMGVDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTO 181
QY 181 VEDCWLGDNVTSLPDLDTTKDVVKQNEWYDMVGSLSVSNYSIDGLRIDTVKHVQKDFWPGYN 240
DB 182 VEDCWLGDNVTSLPDLDTTKDVVKQNEWYDMVGSLSVSNYSIDGLRIDTVKHVQKDFWPGYN 241
QY 241 KAAGVYICIGEVLDGDPAYTCPYQNVMDGVNLPIYYPYLLNAPFKSTSGSMDLLNMINTVK 300
DB 242 KAAGVYICIGEVLDGDPAYTCPYQNVMDGVNLPIYYPYLLNAPFKSTSGSMDLLNMINTVK 301
QY 301 SDCPSTLLGTFTVENHNDNPRFASYNNDIALAKNVAAPFIILNDGIPYIYAGQEQHYAGGND 360
DB 302 SDCPSTLLGTFTVENHNDNPRFASYNNDIALAKNVAAPFIILNDGIPYIYAGQEQHYAGGND 361
QY 361 PANRENTWLSGYPTDSELYKLITASANAIRYVAISKDTGFVYKQWPIYKDDTTIAMRKGT 420
DB 362 PANRENTWLSGYPTDSELYKLITASANAIRYVAISKDTGFVYKQWPIYKDDTTIAMRKGT 421
QY 421 DGSQIVTILSNKAGSDSYTSLSGAGYTAGOOLTEVIGCTTVTGSDGNGVPMAGGLP 480
DB 422 DGSQIVTILSNKAGSDSYTSLSGAGYTAGOOLTEVIGCTTVTGSDGNGVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSSS 498
DB 482 RVLYPTEKLAGSKICSSS 499
RESULT 6
Q96TH4 PRELIMINARY; PRT; 499 AA.
ID Q96TH4;
AC Q96TH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Taka-amylose A (EC 3.2.1.1).
GN Name=amyA;
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI840;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amylolytic gene expression in Aspergillus
RT oryzae.";
RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
CC -! SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AB021876; BAA95703.1; -.
DR PIR; JN0588; JN0588.
DR PIR; S04549; ALAS3.
DR HSP; P10529; 7TAA.
DR GO; GO:0004556; F:alpha-amylose activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_aml_cat.
DR InterPro; IPR006589; Alp_aml_cat_sub.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; Alpha-amylose; 1.
DR PRINTS; PR00110; ALPHAAMYLAASE.

```

```
DR SMART; SM00642; Aamy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match      99.6%; Score 2674; DB 2; Length 499;
Best Local Similarity 99.6%; Pred. No. 1e-177;
Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVAWWSFLYGLQVAAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATCMTADKCYCGG 60
DB 2 MVAWWSFLYGLQVAAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATCMTADKRYCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAYGDAYHGWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAYGDAYHGWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTQ 180
DB 122 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTQ 181
QY 181 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWSGLVSNYSIDGLRIDTVKHVQKDFWPGYN 240
DB 182 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWSGLVSNYSIDGLRIDTVKHVQKDFWPGYN 241
QY 241 KAAGVYCIGEVLGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDLLYNNINTVK 300
DB 242 KAAGVYCIGEVLGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDLLYNNINTVK 301
QY 301 SDPCPSTLLGTFFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGND 360
DB 302 SDPCPSTLLGTFFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGIP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGIP 481
QY 481 RVLYPEKLAGSKICSSS 498
DB 482 RVLYPEKLAGSKICSSS 499

RESULT 7
Q7LV45 PRELIMINARY; PRT; 499 AA.
ID Q7LV45
AC Q7LV45;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Alpha-amylase (EC 3.2.1.1).
GN Name=amy1;
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5059;
[1]
SEQUENCE FROM N.A.
RC STRAIN=86-10D;
RA Fakhoury A.M., Woloshuk C.P.;
RT "Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914 (1999).
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR EMBL; AF139925; AF14264.1; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_ami1 cat.
DR InterPro; IPR006589; Alp_ami1 cat sub.
DR InterPro; IPR006046; Glyco_hydro_13.

DR Pfam; PF00128; Alpha-amylase; 1.
DR PRINTS; PRO0110; ALPHAAMYLASE.
DR SMART; SM00642; Aamy; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match      99.6%; Score 2674; DB 2; Length 499;
Best Local Similarity 99.6%; Pred. No. 1e-177;
Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVAWWSFLYGLQVAAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATCMTADKCYCGG 60
DB 2 MVAWWSFLYGLQVAAAPALAAATPADWRSQSIYFLLTDRFARTDGGSTTATCMTADKRYCGG 61
QY 61 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAYGDAYHGWQDDIYSLNENYGTADDL 120
DB 62 TWQGIIDKLDYIQGMGFTAIWITPVTAQLPOTTAAYGDAYHGWQDDIYSLNENYGTADDL 121
QY 121 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTQ 180
DB 122 KALSSALHERGMYLMDVVVANHMGYDAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQTQ 181
QY 181 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWSGLVSNYSIDGLRIDTVKHVQKDFWPGYN 240
DB 182 VEDCWLGNTVSLPDLDTTKDVVKNEWYDWSGLVSNYSIDGLRIDTVKHVQKDFWPGYN 241
QY 241 KAAGVYCIGEVLGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDLLYNNINTVK 300
DB 242 KAAGVYCIGEVLGDPAYTCPCYQNVMDGVNLPIYYPILLNAFKSTSGSMDLLYNNINTVK 301
QY 301 SDPCPSTLLGTFFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGND 360
DB 302 SDPCPSTLLGTFFVENHNDNPRFASYNNDIALAKNVAAFIILNDGIPIIYAGQSQHYAGGND 361
QY 361 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 420
DB 362 PANREATWLSGYPTDSELYKLIASANAIRNVAISKDTGFTVYKNWPIYKDDTTIAMRKGT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGIP 480
DB 422 DGSQIVTILSNKGASGDSYTLISLGGAGYTAGQQLTEVIGCTTIVTSGDGNVPPVPMAGGIP 481
QY 481 RVLYPEKLAGSKICSSS 498
DB 482 RVLYPEKLAGSKICSSS 499

RESULT 8
AMYA ASPAW STANDARD; PRT; 498 AA.
ID AMYA ASPAW
AC Q02505;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase A).
GN Name=AMYA;
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
[1]
SEQUENCE FROM N.A.
RC STRAIN=UUK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L., Kodama K.H.,
RA Royer T.J., Thompson S.A., Ward M., Wilson L.J., Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212 (1990).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
```

CC at high concentrations (By similarity).
 CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk.
 CC -----
 DR EMBL; X52755; CAA36966.1; --
 DR PIR; A48305; A48305.
 DR HSP; P10529; 7TAA.
 DR InterPro; IPR006589; Alp_amy1_cat_sub.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR InterPro; IPR006046; Glyco_hydro_13.
 DR Pfam; PF00128; Alpha-amy1ase; 1.
 DR PRINTS; PR00110; ALPHAAMYLASE.
 DR SMART; SM00642; Amy; 1.
 KW Calcium-binding; Carbohydrate metabolism; Glycoprotein; Glycosidase;
 KW Hydrolase; Multigene family; Signal.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 498 Alpha-amy1ase A.
 FT ACT_SITE 227 227 Nucleophile (By similarity).
 FT ACT_SITE 251 251 Proton donor (By similarity).
 FT ACT_SITE 318 318 By similarity.
 FT METAL 142 142 Calcium 1 (By similarity).
 FT METAL 183 183 Calcium 1 (via carbonyl oxygen) (By
 FT similarity).
 FT METAL 196 196 Calcium 1 (By similarity).
 FT METAL 227 227 Calcium 2 (By similarity).
 FT METAL 231 231 Calcium 1 (via carbonyl oxygen) (By
 FT similarity).
 FT METAL 251 251 Calcium 2 (By similarity).
 FT DISULFID 51 59 By similarity.
 FT DISULFID 171 185 By similarity.
 FT DISULFID 261 304 By similarity.
 FT DISULFID 461 496 By similarity.
 FT CARBOHYD 218 218 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 498 AA; 54880 MW; 7659511BC01A8A01 CRC64;
 Query Match 99.2%; Score 2662; DB 1; Length 498;
 Best Local Similarity 99.6%; Pred. No. 6.9e-177;
 Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MVAWMSFLYGLQVAPALAAATPADWRSQSIYELLTDRTFARTDGTATCNTADQKYCGG 60
 DB 2 MVAWMSFLYGLQVAPALAAATPADWRSQSIYELLTDRTFARTDGTATCNTADQKYCGG 61
 QY 61 TWQGIIDKLDYIQGMGFTAIWTPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDL 120
 DB 62 TWQGIIDKLDYIQGMGFTAIWTPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDL 121
 QY 121 KALSSALHERGMYLMDVVANHGVDGAGSSVDYSPKPFSSQDYFHPFCFIQNYEDQTO 180
 DB 122 KALSSALHERGMYLMDVVANHGVDGAGSSVDYSPKPFSSQDYFHPFCFIQNYEDQTO 181
 QY 181 VEDCWLGDNTVSLPDLDTTKDVKNENYDWVGSILVSNYSIDGLRIDTVKHVQKDFWPGYN 240
 DB 182 VEDCWLGDNTVSLPDLDTTKDVKNENYDWVGSILVSNYSIDGLRIDTVKHVQKDFWPGYN 241
 QY 241 KAAGVYICIGVLGDGPAYTCYQNVMDGVNLNPIYPLNNAFKSTSGSMDLLNMLNTVK 300
 DB 242 KAAGVYICIGVLGDGPAYTCYQNVMDGVNLNPIYPLNNAFKSTSGSMDLLNMLNTVK 301
 QY 301 SDPCPDSTLLGTFFVENHNDNPFASVTNDIALAKNVAAPFIILNDGIPPIYAGQEQHYAGND 360
 DB 302 SDPCPDSTLLGTFFVENHNDNPFASVTNDIALAKNVAAPFIILNDGIPPIYAGQEQHYAGND 361
 QY 361 PANREATWLSGYTDSLEYKLIASANAIRNVAISKDTGFYTKNWPYKDDTTIARKGT 420
 DB 362 PANREATWLSGYTDSLEYKLIASANAIRNVAISKDTGFYTKNWPYKDDTTIARKGT 421

QY 421 DGSQIVTILSNKGASGDSYTLISGAGYTAGQQLTBIIGCTTIVTGVSDGNVPMAGGLP 480
 DB 422 DGSQIVTILSNKGASGDSYTLISGAGYTAGQQLTBIIGCTTIVTGVSDGNVPMAGGLP 481
 QY 481 RVLYPTEKLAGSKIC 495
 DB 482 RVLYPTEKLAGSKIC 496
 RESULT 9
 QYU07
 ID QYU07 -PRELIMINARY; PRT; 490 AA.
 AC QYU07;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Alpha-amy1ase AmyA.
 GN Name=amyA;
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.M., Kelly J.M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
 DR EMBL; AF208225; AAF17103.1; --
 DR HSP; P10529; 7TAA.
 DR GO; GO:0004556; F:alpha-amy1ase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR InterPro; IPR006589; Alp_amy1_cat_sub.
 DR InterPro; IPR006046; Glyco_hydro_13.
 DR Pfam; PF00128; Alpha-amy1ase; 1.
 DR PRINTS; PR00110; ALPHAAMYLASE.
 DR SMART; SM00642; Amy; 1.
 SQ SEQUENCE 490 AA; 54249 MW; A891C4ACEABE5305 CRC64;
 Query Match 68.0%; Score 1826; DB 2; Length 490;
 Best Local Similarity 69.6%; Pred. No. 1e-118;
 Matches 337; Conservative 53; Mismatches 92; Indels 2; Gaps 2;
 QY 12 LQVAPALAAATPADWRSQSIYELLTDRTFARTDGTATCNTADQKYCGTWOIIGLDY 71
 DB 7 LQFAATVLAATPAWRSQSIYELLTDRTFARTDGTATCNTADQKYCGTWOIIGLDY 65
 QY 72 IQGMGFTAIWTPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERG 131
 DB 66 IQGMGFTAIWTPVTAQLPQTAYGDVHYGQQDIYSLNENYGTADDLKALSSALHERG 124
 QY 132 MYLMDVVANHGVDGAGSSVDYSPKPFSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTV 191
 DB 125 MYLMDVVANHGVDGAGSSVDYSPKPFSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTV 184
 QY 192 SLPLDLTTKDVKNENYDWVGSILVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYICIGV 251
 DB 185 SLPLDLTTKDVKNENYDWVGSILVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYICIGV 244
 QY 252 LQGDPAATCPYQNVMDGVNLNPIYPLNNAFKSTSGSMDLLNMLNTVKSDCPDSTLLGT 311
 DB 245 LQGDPAATCPYQNVMDGVNLNPIYPLNNAFKSTSGSMDLLNMLNTVKSDCPDSTLLGT 304
 QY 312 FVENHNDNPFASVTNDIALAKNVAAPFIILNDGIPPIYAGQEQHYAGNDPANREATWLSG 371
 DB 305 FVENHNDNPFASVTNDIALAKNVAAPFIILNDGIPPIYAGQEQHYAGNDPANREATWLSG 364
 QY 372 YPTDSELYKLIASANAIRNVAISKDTGFYTKNWPYKDDTTIARKGTGDSQIVTILSN 431
 DB 365 YPTDSELYKLIASANAIRNVAISKDTGFYTKNWPYKDDTTIARKGTGDSQIVTILSN 424
 QY 432 KGASGDSYTLISGAGYTAGQQLTBIIGCTTIVTGVSDGNVPMAGGLPRVLYPTEKLAG 491

```
Db 425 AGADAGSSTVSPNTGTTAGAAVTEITCEDITVSGSGSVPMBSGLPRVLVYPKALRG 484
QY 492 SKIC 495
Db 485 SGIC 488
RESULT 10
ID O13296 PRELIMINARY; PRT; 640 AA.
AC O13296;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Acid-stable alpha-amylase.
OS Aspergillus kawachi (Aspergillus awamori var. kawachi).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=40384;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaneko A., Sudo S., Sakamoto Y., Tamura G., Ishikawa T., Ohba T.;
RT "Molecular-cloning and determination of the nucleotide-sequence of a
RT gene encoding an acid-stable alpha-amylase from Aspergillus-
RT kawachi.";
RL J. Ferment. Bioeng. 81:292-298 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Kaneko A.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008370; BAA22993.1; -.
DR HSSP; P56271; 2AAA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR002044; Glyco_hydro_CBD.
DR Pfam; PF00128; Alpha-amylase; 1.
DR Pfam; PF00686; CBM_20; 1.
DR ProDom; PD001568; Glyco_hydro_CBD; 1.
DR SMART; SM00642; Amy; 1.
SQ SEQUENCE 640 AA; 69507 MW; 676BF9D0236720DD CRC64;

Query Match 67.4%; Score 1810; DB 2; Length 640;
Best Local Similarity 67.6%; Pred. No. 1.9e-117;
Matches 325; Conservative 66; Mismatches 90; Indels 0; Gaps 0;

QY 18 ALAATPADWRSSQSIYFLLTDRFARTDGTATCNTADQKCGTGWQGIIDKLDYIQGMGF 77
Db 19 ALGLSAAEWRTQSIYFLLTDRFRTDNTATCNTGDIYCGSGWQGIINHLLDIQGMGF 78

QY 78 TAIWTPVTAQLPQTAYGDVGHVWQDIYSLNENYGTADDLKALSSALHGRGMYLMD 137
Db 79 TAIWISPITEQLPQDTSDEAYHGVWQDIYVNSNFGTADDLKSLSDALHARGMYLMD 138

QY 138 VVANHMGYDGAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTVEDCWLDGNTVSLPDLN 197
Db 139 VVPNHMGYAGNGNDVDSVDFPDSYFHPYCLITDNDNLTMVQDCWEGDTIVSLPDLN 198

QY 198 TTKDVVKNEWDVWVGLSVNSYSDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDA 257
Db 199 TTETAVRTIWDVADLVNSYSDGLRIDSVLEVEFPDFPGYQEAAGVYCVGEVDNGPA 258

QY 258 YTCPCYQNVMDGVNLPIYYPYLLNAPFKSTSGMDDLNNMTVKSCDPCDSTLLGTFFVENHD 317
Db 259 LDCPYOKYLDGVNLPIYVWQLLYAFESSSGSISLDYNNIKSVASDCSDPTLLGNFNIENHD 318

QY 318 NPREASYNTDIALAKNVAAFIILNDGIPITYAGOEQHYAGGNDPANREATWLSGYPTDSE 377
Db 319 NPREFASYSYDQAKNVLSYIFLSDGIPITYAGEEQHYSGGDPVFNREATWLSGYDTSAE 378
```

```
QY 378 LYKLIASANAIRNVAISKDTGFTVYKNPIYKDDTTIAMRKCTDGSQIVTILSNKGASGD 437
Db 379 LYTWIATTAIRKLAISADSDVITYKNDPIYTDSTNTIAMRKCTSGSQIITVLSNKGSSGS 438

QY 438 SYTSLSGAGYTAGOOLTEVIGCTTIVTSGDGNVPMAGGLPRVLYPTEKLAGSKICSS 497
Db 439 SYTSLSGSGYTSGLKLEIAYTCTSVTDSNGDIPVPMASGLPRVLLPASVVDSSLCGG 498

QY 498 S 498
Db 499 S 499

RESULT 11
QY 76L96 PRELIMINARY; PRT; 634 AA.
ID Q76L96
AC Q76L96;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Alpha-amylase precursor.
GN Name=amyl III.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT-11;
RA Matsubara T.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083162; BAD06005.1; -.
DR EMBL; AB083160; BAD06003.1; -.
DR HSSP; P04064; 1ACZ.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR002044; Glyco_hydro_CBD.
DR Pfam; PF00128; Alpha-amylase; 1.
DR Pfam; PF00686; CBM_20; 1.
DR ProDom; PD001568; Glyco_hydro_CBD; 1.
DR SMART; SM00642; Amy; 1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 634 AA; 69242 MW; 16C0B6AF6F0E9B CRC64;

Query Match 66.7%; Score 1791; DB 2; Length 634;
Best Local Similarity 68.6%; Pred. No. 3.9e-116;
Matches 321; Conservative 64; Mismatches 83; Indels 0; Gaps 0;

QY 18 ALAATPADWRSSQSIYFLLTDRFARTDGTATCNTADQKCGTGWQGIIDKLDYIQGMGF 77
Db 19 ALGLSAAEWRTQSIYFLLTDRFRTDNTATCNTGDIYCGSGWQGIINHLLDIQGMGF 78

QY 78 TAIWTPVTAQLPQTAYGDVGHVWQDIYSLNENYGTADDLKALSSALHGRGMYLMD 137
Db 79 TAIWISPITEQLPQDTSDEAYHGVWQDIYVNSNFGTADDLKSLSDALHARGMYLMD 138

QY 138 VVANHMGYDGAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTVEDCWLDGNTVSLPDLN 197
Db 139 VVPNHMGYAGNGNDVDSVDFPDSYFHPYCLITDNDNLTMVQDCWEGDTIVSLPDLN 198

QY 198 TTKDVVKNEWDVWVGLSVNSYSDGLRIDTVKHVKQDFWPGYNKAAGVYCIGEVLDGDA 257
Db 199 TTETAVRTIWDVADLVNSYSDGLRIDSVLEVEFPDFPGYQEAAGVYCVGEVDNGPA 258

QY 258 YTCPCYQNVMDGVNLPIYYPYLLNAPFKSTSGMDDLNNMTVKSCDPCDSTLLGTFFVENHD 317
Db 259 LDCPYQDYLDGVNLPIYVWQLLYAFESSSGSISLDYNNIKSVASDCSDPTLLGNFNIENHD 318

QY 318 NPREASYNTDIALAKNVAAFIILNDGIPITYAGOEQHYAGGNDPANREATWLSGYPTDSE 377
Db 319 NPREFASYSYDQAKNVLSYIFLSDGIPITYAGEEQHYSGGNDPANREATWLSGYPTDSE 377
```

Db 319 NRPASVTSYQAKWLSVIFLSDGPIVAGEBHYSGDVPYREATWLSGYDTSAE 378
QY 378 LYKLIASAINAIRNAYISKOTGFVYKXNPIYKDDTTIAMRKGTGDSQIVTILSNKGASGD 437
Db 379 LYTWIATTAIRKLAISADSDYITANDPIYDTSNTIAMRKGTSGQVITVLSNKGSSGS 438
QY 438 SYTSLSGAGYTAGQQLTEVIGCTTIVTSGDGNVPVPMAGGLPRVLXP 485
Db 439 SYTLTSLSGGYTSGTELEIAYTCTSVTDSNGDIPVPMASGLPRVLLP 486

RESULT 12
MYA_ASPNG STANDARD; PRT; 484 AA.
AC P56271;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Acid alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
OS *Aspergillus niger*.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
OX NCBI_TaxID=5061;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=91002514; PubMed=2207069;
RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
RA Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.;
RT "Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-
RT A resolution of two enzymes from *Aspergillus*.";
RL Biochemistry 29:6244-6249(1990).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -I- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
CC at high concentrations.
CC -I- SUBUNIT: Monomer.
CC -I- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.
DR PDB; 2AAA; X-ray; @=1-484.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR Pfam; PF00128; Alpha-amylase; 1.
DR SMART; SM00642; Amy; 1.
KW 3D-structure; Calcium-binding; Carbohydrate metabolism; Glycoprotein;
KW Glycosidase; Hydrolase.
FT ACT_SITE 206 206 Nucleophile.
FT ACT_SITE 230 230 Proton donor.
FT ACT_SITE 297 297
FT METAL 121 121 Calcium 1.
FT METAL 162 162 Calcium 1 (via carbonyl oxygen).
FT METAL 175 175 Calcium 1.
FT METAL 206 206 Calcium 2.
FT METAL 210 210 Calcium 1 (via carbonyl oxygen).
FT METAL 230 230 Calcium 2.
FT DISULFID 30 38
FT DISULFID 150 164
FT DISULFID 240 283
FT DISULFID 440 475
FT CARBOHYD 24 24
FT CARBOHYD 157 157 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 197 197 N-linked (GlcNAc...) (Potential).
FT HELIX 3 6
FT TURN 7 8
FT STRAND 11 13
FT HELIX 15 19
FT TURN 22 23
FT HELIX 32 34
FT TURN 42 47
FT TURN 48 48
FT HELIX 49 53
FT TURN 54 56
FT STRAND 59 62
FT STRAND 66 68

Query Match

484 AA; 52935 MW; 04D596E34680656D CRC64;

66.2%; Score 1778; DB 1; Length 484;

Matches		298;	Conservative	75;	Mismatches	110;	Indels	2;	Gaps	1;
QY	12	LQVAPALAAATPADWRSQSIYFLLTRFARTDGTATNTADQKCYGGTWQGIIDKLDY	71							
Db	164	ISVSVP--TCTAANWRGRAIYQIVTRFARTDGTSTTCLDVTDRVYCGSGYQGIINMLDY	221							
QY	72	IQMGFTAIWTPVTAQLPQTAYGDAYHCYWOODIYSLNENYGTADDLKALSSALHRS	131							
Db	222	IQMGFTAIWISPIVENIPDDTGYGAYHYGWMKDIFALNTNFGTADDLIALATELHNRG	281							
QY	132	MYLMDVVAANMGYDAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTV	191							
Db	282	MYLMDVIVNHFASFSGHADVDYSEYFPYSSQDYFHSFCWITDYSNQTVNVEQCMLGDDIV	341							
QY	192	SLPDLTTKDVVKNWYDVGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYICIGEV	251							
Db	342	PLVDVNTQDITVKSEYQSWQVQLIANYSIDGLRIDTVKHVQKDFWPAFQEAAGIYAVGEV	401							
QY	252	LDGDPATCPYQNVMDGVLPYIYPLNAFKSTSGSMDDLYNMINTVKSDCPDSTLLGT	311							
Db	402	FDGDSYTCPYQENLDGVLPYIYVWVSFAFESVSGSVSLVMDITLSECTDITLLGS	461							
QY	312	FVENHNPFPASTNDIALAKNVAAFIILNDGPIIYAGOEQHYAGNDPANREATWLSG	371							
Db	462	FLENQNPFPSPSTSDSLKNAIAFTMLSDGPIIYGOEQLNGNDPNYREALWLIG	521							
QY	372	YPTDSLYKLIAANAIYNSKDTGFTVYKNWPIYKDDTTIAMRGTDGSGIVTILSN	431							
Db	522	YSTTSFYKYIASLQIRNAIQKDDTYLTQYQNVWISDSTTIAMKFTGNQIITVLSN	581							
QY	432	KGASGDSYTLISGAGYTAQQQTEVIGCTTVTVGSDGNVPVPMAGLPRVLYPTBKL	491							
Db	582	LGTSGSSYTLTSLNTGYTASSVVEILTCTAVTVDSGSLAVPMWGLPKRVFYQESQ	641							
QY	492	SKICS	496							
Db	642	SGICS	646							
RESULT 15										
AMV1_LIPKO STANDARD; PRT; 624 AA.										
AC	Q0117;									
DT	10-OCT-2003 (Rel. 42, Created)									
DT	10-OCT-2003 (Rel. 42, Last sequence update)									
DT	25-OCT-2004 (Rel. 45, Last annotation update)									
DE	Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan									
DE	glucanohydrolase 1).									
GN	Names=LKAI;									
OS	Lipomyces kononenkoae.									
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;									
OC	Saccharomycetales; Lipomycetaceae; Lipomyces.									
OX	NCBI_TaxID=34357;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=IGC4052B;									
RX	MEDLINE=96105202; PubMed=8529895; DOI=10.1016/0378-1119(95)00633-0;									
RA	Steyn A.J.C., Marmur J., Precorius I.S.;									
RT	"Cloning, sequence analysis and expression in yeasts of a cDNA									
RT	containing a Lipomyces kononenkoae alpha-amylase-encoding gene.";									
RL	Gene 166:65-71 (1995).									
RN	[2]									
RP	SEQUENCE OF 29-44.									
RC	STRAIN=IGC4052B;									
RX	MEDLINE=96132108; PubMed=8593683;									
RA	Steyn A.J.C., Precorius I.S.;									
RT	"Characterization of a novel alpha-amylase from Lipomyces kononenkoae									
RT	and expression of its gene (LKAI) in Saccharomyces cerevisiae.";									
RL	Curr. Genet. 28:526-533 (1995).									
CC	!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic									
CC	linkages in oligosaccharides and polysaccharides.									
CC	!- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory									
CC	at high concentrations (By similarity).									

!- SUBCELLULAR LOCATION: Secreted.										
!- SIMILARITY: Belongs to the glycosyl hydrolase 13 family.										
!- SIMILARITY: Contains 1 carbohydrate binding type-21 (CBM21) domain.										

This SWISS-PROT entry is copyright. It is produced through a collaboration										
the European Bioinformatics Institute. There are no restrictions on its										
use by non-profit institutions as long as its content is in no way										
modified and this statement is not removed. Usage by and for commercial										
entities requires a license agreement (See http://www.isb-sib.ch/announce/										
or send an email to license@isb-sib.ch).										

CC	EMBL; U30376; AAC49622.1; ALT_INIT.									
DR	PIR; JC4510; JC4510.									
DR	HSSP; P10529; 7TRA.									
DR	InterPro; IPR006589; Alp amyl cat sub.									
DR	InterPro; IPR006047; Alpha amyl cat.									
DR	InterPro; IPR005036; CBM 21.									
DR	Pfam; PF00128; Alpha-amylase; 1.									
DR	Pfam; PF03370; CBM 21; 1.									
DR	SMART; SM00642; Amy; 1.									
DR	Calcium-binding; Carbohydrate metabolism; Direct protein sequencing;									
KW	Glycoprotein; Glycosidase; Hydrolase; Signal.									
FT	SIGNAL 1 28									
FT	CHAIN 29 624									
FT	DOMAIN 40 133									
FT	ACT_SITE 353 353									
FT	ACT_SITE 377 377									
FT	ACT_SITE 444 444									
FT	METAL 268 268									
FT	METAL 309 309									
FT	METAL 322 322									
FT	METAL 353 353									
FT	METAL 357 357									
FT	METAL 377 377									
FT	DISULFID 177 185									
FT	DISULFID 297 311									
FT	DISULFID 387 430									
FT	DISULFID 587 622									
FT	CARBOHYD 304 304									
FT	CARBOHYD 344 344									
SEQ	SEQUENCE 624 AA; 68876 MW; 87EB16534F5A9A9F CRC64;									
Query Match 60.7%; Score 1629; DB 1; Length 624;										
Best Local Similarity 60.6%; Pred No. 7.1e-105;										
Matches 292; Conservative 77; Mismatches 113; Indels 0; Gaps 0;										
QY	15	AAPALAAATPADWRSQSIYFLLTRFARTDGTATNTADQKCYGGTWQGIIDKLDYIQG	74							
Db	142	SASVPTGTAANWRGSIYQVTVTRFARTDGTATNTADQKCYGGTWQGIIDKLDYIQG	201							
QY	75	MGFTAIWTPVTAQLPQTAYGDAYHCYWOQDIYSLNENYGTADDLKALSSALHRS	134							
Db	202	MGFTAIWISPIVENIPDDTGYGAYHYGWMKDIFALNTNFGTADDLIALATELHNR	261							
QY	135	MYDVVVAANMGYDAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTV	194							
Db	262	MYDVVVAANMGYDAGSSVDYSVFKPSSQDYFHPFCFIQNYEDQTOVEDCWLGDNTV	321							
QY	195	DLDTTKDVVKNWYDVGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAAGVYICIGEV	254							
Db	322	DVNTQDITVKSEYQSWQVQLIANYSIDGLRIDTVKHVQKDFWPAFQEAAGIYAVGE	381							
QY	255	DPATCPYQNVMDGVLPYIYPLNAFKSTSGSMDDLYNMINTVKSDCPDSTLLGT	314							
Db	382	DPSYTCPYQENLDGVLPYIYVWVSFAFESVSGSVSLVMDITLSECTDITLLGS	441							
QY	315	NHDNPRFASVTNDIALAKNVAAFIILNDGPIIYAGOEQHYAGNDPANREATWLSG	374							
Db	442	NQDNPRFSPSTSDSLKNAIAFTMLSDGPIIYGOEQLNGNDPNYREALWLIG	501							

QY	375	DSELYKLIASANAIRNYAISKDTGFTYKXWPIYKDDTTIAMRKGTGDSQIVTILSNKGA	434
Db	502	TSTFVEYIASLNQIRNHAIYIDDTVLTQNWVIYSDSTTIAMRKGTGNQIITVLSNLGS	561
QY	435	SGDSYTLISLGGAGYTAGOOLTEVIGCTTVTVGSDGNVPVPMAGGLPRVLYPTEKLAGSKI	494
Db	562	SGSSYTLTILSNLTGYTASSVVEILTCTAVTVDLGNLAVPMMSGGLPRVFYPSQLVSGSI	621
QY	495	CS	496
Db	622	CS	623

Search completed: February 8, 2005, 18:41:52
Job time : 300.263 secs